A Data-driven Approach for Identifying Ethical Concerns of Climate Engineering Technologies - Stratospheric Aerosol Injection as a use-case

Julius Fenn[1], Michael Gorki[1, 2], Philipp Höfele[3], Louisa Estadieu[2,4], Christophe Becht[1], Lars Kulbe[1, 2], Johannes Gekeler[1] and Andrea Kiesel[1, 2]

[1]Institute of Psychology, University of Freiburg, Germany

[2]Cluster of Excellence livMatS @ FIT Freiburg Center for Interactive Materials and Bioinspired Technologies, University of Freiburg, Germany

[3]Young Academy for Sustainability Research (YAS) at Freiburg Institute for Advanced Studies (FRIAS), University of Freiburg, Germany

[4] Department of Humanities, Social and Political Sciences, ETH Zürich, Switzerland

Author Notes

Julius Fenn https://orcid.org/0000-0002-7569-8648

E-Mail: julius.fenn@psychologie.uni-freiburg.de

Michael Gorki https://orcid.org/0009-0007-6138-4043

E-Mail: michael.gorki@livmats.uni-freiburg.de

Philipp Höfele https://orcid.org/0000-0002-8682-9965

E-Mail: philipp.hoefele@frias.uni-freiburg.de

Louisa Estadieu https://orcid.org/0000-0002-3164-4445

E-Mail: louisa.estadieu@ethz.ch

Christophe Becht https://orcid.org/0009-0005-8483-2588

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=5272391

E-Mail: christophe.becht@psychologie.uni-freiburg.de

Lars Kulbe https://orcid.org/0009-0005-6244-6002

E-Mail: lars.kulbe@student.uva.nl

Johannes Gekeler https://orcid.org/0009-0007-4936-5632

E-Mail: johannes.gekeler@students.uni-freiburg.de

Andrea Kiesel https://orcid.org/0000-0001-5564-010X

E-Mail: kiesel@psychologie.uni-freiburg.de

Lars Kulbe is now at the Department of Psychology, University of Amsterdam.

*corresponding author: Julius Fenn, Department of Psychology, University of Freiburg, Engelbergerstraße 41, 79085 Freiburg. julius.fenn@psychologie.uni-freiburg.de

Funding

The work leading to this publication was supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy – EXC-2193/1 – 390951807, as well as by the Eva Mayr-Stihl Foundation as part of the "Young Academy for Sustainability Research" at the "Freiburg Institute for Advanced Studies" (FRIAS), University of Freiburg.

Acknowledgement

This work was inspired by discussions with our colleagues from the Living, Adaptive and Energy-autonomous Materials Systems excellence cluster of the University of Freiburg. Many thanks to my co-author Philipp Höfele for making me aware of the importance of an empirically informed ethics.

Declaration of Interest statement

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=5272391

Abstract

Limiting global warming to 1.5 °C has intensified interest in climate engineering technologies such as Stratospheric Aerosol Injection (SAI), which mimic volcanic cooling. Given potential insufficiency of mitigation alone, ethical examination of SAI is imperative. This study investigates whether laypersons' ethical reasoning about SAI can be empirically identified. Using a multi-method design, we combined Cognitive-Affective Maps (CAMs) and open-ended textual responses to elicit twenty ethical concerns. Large Language Models (LLMs) synthesized lay perspectives and compared them against formal ethical definitions. Results revealed diverse ethical considerations, including governance, risk, equitable deployment, and emergency use. In contrast to formal definitions, lay participants foregrounded practical implications, social trust, and personal experience. Our findings demonstrate the utility of integrating data sources for empirical ethics research and underscore the complexity of public ethical discourse on SAI. This approach promotes more inclusive, evidence-based dialogue on the responsible development and governance of climate engineering technologies.

Keywords: climate engineering; climate change; Cognitive-Affective Maps; qualitative content analysis; network analysis; large language models

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=5272391

1. Introduction

Current attempts to lower greenhouse gas emissions and to adapt to the effects of climate change are still insufficient according to multiple authors (e.g., Le Quéré et al., 2021; Lee & Romero, 2023; Pörtner et al., 2022; Welsby et al., 2021) or as indicated by recent reports of the Climate Action Tracker[1]. The most recent report of the Intergovernmental Panel on Climate Change (IPCC, AR6 Synthesis Report from 2023) warns that current mitigation contributions “make it likely that warming will exceed 1.5°C during the 21st century” (Lee & Romero, 2023, p. 23). This concern is underscored by the fact that global average temperatures surpassed the 1.5°C threshold above pre-industrial levels for the first time in 2024[2]. In the perspective of the environmental philosopher Gardiner (Gardiner, 2011) this marks an “environmental tragedy” – despite (scientific) facts on climate change being well known (e.g., in The Limits to Growth report by Meadows et al., 1972), current actions are not sufficiently effective.

To increase the chances that temperature increase is limited to 1.5°C, multiple reports and scientific studies emphasize that Climate Engineering Technologies (CETs), especially “negative emissions” technologies, are necessary (e.g., Anderson & Peters, 2016; Haszeldine et al., 2018; Johansson et al., 2020; Welsby et al., 2021). In general, there are two distinct approaches of CETs to address climate change (see Caviezel & Revermann, 2014; Dowling, 2018; Heyward, 2013; National Research Council, 2015; Shepherd, 2009) with varying ethical concerns (see Betz & Cacean, 2012; Ott & Neuber, 2020; Rickels et al., 2011). Carbon Dioxide Removal technologies, also called “negative emissions” technologies,

---

[1] Climate Action Tracker is an independent scientific project that analyzes data from 39 countries, collectively covering 85% of global emissions, to produce its reports; see Climate Action Tracker Thermometer: https://climateactiontracker.org/global/cat-thermometer/

[2] See “Copernicus: 2024 is the first year to exceed 1.5°C above pre-industrial level”: https://climate.copernicus.eu/copernicus-2024-first-year-exceed-15degc-above-pre-industrial-level; The Copernicus Climate Change Service (C3S), operated by the European Centre for Medium-Range Weather Forecasts (ECMWF) as part of the EU's Copernicus programme, provides free, reliable, and up-to-date data on climate and environmental changes.

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=5272391

remove carbon dioxide from the atmosphere, which addresses the root cause of climate change. In contrast, Solar Radiation Management technologies seek to reflect a small percentage of solar radiation back into space before it reaches the earth. Such technologies are already included in most scenarios (so-called "Integrated Assessment Models") of the IPCC reports, which quantitatively describe key human and earth system processes of climate change (e.g., Lee & Romero, 2023; Masson-Delmotte et al., 2018; Pachauri & Meyer, 2014; Pörtner et al., 2022).

According to Sand et al. (2023), CETs can be framed as a "techno-fix" for the problem of insufficient climate mitigation. Such CETs do not demand behavioral changes of people and might be implemented more easily and faster than large societal transformations (Preston, 2012, 2013). Because such technologies could free ourselves from the obligation to reduce emissions and thereby impact our moral agency (Gardiner, 2010a), framing CETs as a "techno-fix" is therefore highly contested from an ethical standpoint (Corner & Pidgeon, 2014). Due to, for example, possible unknown side-effects, the problem of climate change could even be enlarged (see expert interviews in Sovacool et al., 2022, 2023).

Given the critical role of CETs in addressing climate issues, it is critical to empirically investigate their ethical concerns associated with their development and implementation. Such inquiry could finally support the responsible and informed governance of these emerging technologies (Low et al., 2024; Reynolds & Horton, 2020). To this end, we focus on a specific Solar Radiation Management technology known as Stratospheric Aerosol Injection (SAI) as a use case, yet our proposed methodology can be easily adjusted and applied to different types of emerging CETs. Investigating ethical concerns of SAI is crucial because the technology is highly efficient in comparison to other CETs (D. P. Keller et al., 2014; Sonntag et al., 2018), timely and relatively cheap (Shepherd, 2009). SAI can decrease the amount of incoming solar radiation by releasing sulfur particles into the stratosphere,

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=5272391

enhancing the aerosol layer's reflective properties. This technology, most prominently proposed by Crutzen (2006), mimics the natural cooling effect observed after volcanic eruptions (e.g., Mount Tambora in 1815 or Mount Pinatubo in 1991), during which sulfur particles are released into the atmosphere (cf., Plazzotta et al., 2018; Zhang et al., 2022). SAI could be deployed in an emergency case when mitigation efforts have been insufficient. However, there are fundamental ethical concerns (for an overview, see Tab. 1 in section 1.2.), which lead some scientists to advocate a Non-Use Agreement (Biermann et al., 2022). For example, even the act of just researching SAI could by itself decrease the motivation of individuals and governments to implement necessary, far-reaching mitigation policies (ethical argument of "Moral Hazard") and transfer the risks of climate change to future generations, thus putting them in a dilemma to finally deploy SRM technologies, which is the ethical argument of "Risk Transfer to the Future" (e.g., Callies, 2019; Preston, 2012).

We propose a methodology that relates two heterogeneous data sources - Cognitive-Affective Maps (CAMs) and open text - and apply three types of data analyses: network analyses, qualitative content analysis, and Large Language Models (LLMs) to answer our main research question: Is it possible to empirically identify ethical arguments of laypersons regarding our use-case SAI? The article is organized as follows: In Section 1.1, we briefly motivate the need for empirically informed ethics, followed by a discussion of individual ethical arguments in Section 1.2. Section 2 describes the overall study design, which includes two time points for data collection and two different types of data. Section 3 presents the results for both data sources, along with their respective statistical procedures. Finally, Section 4 provides an overview of all results and proposes future research questions.

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=5272391

1.1. Motivation of empirical informed ethics

The 15th Principle of the Rio Declaration on Environment and Development, the so-called Precautionary Principle, states that if “there are threats of serious or irreversible damage, lack of full scientific certainty shall not be used as a reason for postponing cost-effective measures to prevent environmental degradation” (United Nations Environment Programme, 1992, p. 6). Thus, it could be argued that CETs, which are at an early development stage, should be evaluated regarding benefits and costs while also considering risks which are not of “full scientific certainty” (cf., Buckley et al., 2017). To find an optimal balance between the economic costs of greenhouse gas reductions (e.g., reduced consumption) and their benefits (e.g., reduced gross damages) based on the Dynamic Integrated Climate-Economy Model by Nordhaus (1992), multiple climate scenarios have been simulated (e.g., De Bruin et al., 2009; Johansson et al., 2020; Nordhaus, 2018). There is a vivid discussion in the economic literature, for instance, on whether CETs are relatively inexpensive compared to mitigation (see Barrett, 2008 vs. Klepper & Rickels, 2012) or on how the parameters of such models should be adjusted to account for factors such as intergenerational welfare (Hänsel et al., 2020).

Relying exclusively on expert-driven strategies or climate scenario modeling for decision-making in addressing climate change is problematic due to deep uncertainties inherent in the dynamics of the climate system and the complexity of the planetary boundaries we are approaching (e.g., Rockström et al., 2009). Neither is there expert consensus on what outcomes (e.g., cost-effectiveness vs. intergenerational equity) should be aimed for, nor which climate policies should be pursued (K. Keller et al., 2021; Marchau et al., 2019; Workman et al., 2020). For CETs specifically, irreducible uncertainties persist, such as those arising from the inherent complexities of the Earth system, human error, and limitations in predictive models (e.g., Betz & Cacean, 2012; Neuber, 2018; Rickels et al.,

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=5272391

2011), underscoring the need for a nuanced and multi-perspective approach to their evaluation. We argue that it is at least central to ethically evaluate emerging CETs in real time to decide between a cautious (more conservative) approach, linked to the principle of precaution (cf., Höfele, 2020; Jonas, 2020), and a constructive (liberal) approach, linked to an optimistic principle of innovation (cf. Bindé, 2000; Grunwald, 2014; Guston & Sarewitz, 2002; Musschenga, 2009).

To guide decision-making processes under conditions of such high system uncertainty and high decision stakes, Funtowicz and Ravetz (2018) emphasize the necessity of including "extended peer communities". Standard expert-analytical assessments alone are inadequate in such contexts, which are emblematic of "post-normal science" (cf., Workman et al., 2020), where the complexity and stakes of decisions demand broader participation and diverse perspectives. Incorporating laypersons' perspectives alongside expert-driven approaches not only makes decisions more attuned to diverse values and ethical concerns, generating new substantive insights, but also enhances the legitimacy of decisions and fosters greater trust in policymaking processes (Fiorino, 1990; Pidgeon, 2021; Wibeck et al., 2017).

The inclusion of laypersons could even foster structures of *anticipatory governance*, which is the capacity "extended through society that can act on a variety of inputs to manage emerging knowledge-based technologies while such management is still possible" (Guston, 2014, p. 219). It is also possible, for example by using robust decision making, to identify climate policy options which are robust in many possible future scenarios, whereby the process of making decisions adapts to the unfolding future (Marchau et al., 2019). In the context of CETs, considering the (ethical) concerns of all stakeholders affected by such technologies would, at best, lead to changes in the research and implementation process (cf., Frumhoff & Stephens, 2018; Gardiner & Fragnière, 2018). In our opinion, such an empirically informed ethics, i.e. the integration of laypersons' perspectives at the early stages

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=5272391

of CET development, could enable the identification of potential ethical challenges and foster adaptive, socially responsive approaches to governance.

In the next section, we present an overview of ethical arguments regarding CETs already identified in the literature. These ethical arguments are applied to structure the view of laypersons and relate their answers (see CAMs, open text sections) to ethically established arguments (for a similar procedure see Höfele et al., 2022).

1.2. Ethical Arguments

CETs and emerging technologies in general are accompanied by uncertainties that call for normative regulations. Thereby, the appropriate use and design of technologies, as well as the acceptable consequences associated with them, are unknown (cf., Grunwald, 2004, 2022). Such uncertainties are especially dealt with by means of two interrelated ethical perspectives (see Cotton, 2014; Grunwald & Hillerbrand, 2021; Pieper, 2017). *Normative ethics* investigates the criteria for determining the moral rightness or wrongness of actions and virtues. *Applied ethics* involves applying ethical principles or theories to specific problems and conflicts in various life areas. It has developed into several independent subfields, including, for example, medical, environmental, animal, science and technology, political, legal, professional, and business ethics, which have expanded significantly over the past 20 years (Grunwald & Hillerbrand, 2021; Neuhäuser et al., 2023). *Empirical ethics* goes a step further by integrating empirical data to examine how moral values are understood and lived in practice, enriching the normative analysis with insights from real-world behaviors and attitudes (Paulo & Bublitz, 2020). Ethical arguments related to SAI belonging to the realm of applied ethics often take a deductive form (argument is deductively valid if its conclusion logically follows from the premises) and make use of descriptive empirical and normative premises (cf., Betz & Cacean, 2012; Neuber, 2018). For example, the "Lesser-Evil" argument

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=5272391

states that deploying SAI is necessary to prevent catastrophic global warming, and includes a descriptive premise ("[a]t some future point in time t, we may end up in a situation where [...] the worst possible impacts of the deployment of the CET are clearly less severe than the worst possible consequences of not deploying it"; Betz & Cacean, 2012, p. 32) and a normative premise ("one should choose the option for action with the comparatively best worst possible consequences"; Betz & Cacean, 2012, p. 32). Due to the reliance on often changing descriptive premises and their complexities, ethical arguments often cannot be definitively justified as true or right, in our opinion. Therefore, they need to be continually critically evaluated to identify the best evidence (descriptive premises) for a certain conclusion (e.g., deploying SAI is the best option in a specific future context). Such a procedure is closely linked to the theory of "The Inference to the Best Explanation" (Harman, 1965; McCain & Poston, 2017), as such an ethical argument "includes relevant considerations that give us reason for thinking that the conclusion is likely to follow" (McMillan, 2018, p. 113).

The ethical arguments regarding CETs in Table 1, which are applied in the following sections, are based on multiple reports and scientific articles from authors in the field of philosophy and ethics (cf., Betz & Cacean, 2012; Neuber, 2018; Ott, 2011, 2012; Ott & Neuber, 2020; Preston, 2012, 2013; Rickels et al., 2011) as well as from authors in the field of the social sciences, whereby we only considered qualitative studies investigating the ethical concerns of laypersons (Carr & Yung, 2018; Corner et al., 2011, 2013; McLaren et al., 2016; Parkhill et al., 2013; Parkhill & Pidgeon, 2011; Pidgeon et al., 2013; Wibeck et al., 2017). We iteratively generated and adjusted definitions of the ethical arguments during team discussions.

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=5272391

Table 1

Overview of identified ethical arguments regarding CETs

Ethical Argument	Definition	Coding Rules	General Evaluation
Moral Hazard (also: “Undermining Better Options”)	• Researching and developing CETs may foster the idea of a technical climate solution, which might reduce people's enthusiasm for pursuing (potentially challenging) mitigation measures / mitigation policies • Solely investing in CETs research and development may divert resources from mitigation efforts • Lobby groups and media hype around CETs could further undermine emission abatement and adaptation measures	Compared to “Arming the Future” negative future perspective on the research of CETs. Compared to “Risk Transfer to the Future” the argument focuses on mitigation efforts / policies (no global perspective). Respondents do not need to emphasize the last bullet point.	Negative
Risk Transfer to the Future	• Research and development of CETs transfers risks to future generations • CETs can create new conflicts and may trigger wars • Deciding to deploy or not deploy these technologies will likely lead to future dilemmas	Compared to “Arming the Future” negative future perspective on the research of CETs. Compared to “Moral Hazard” this argument takes a more global perspective (e.g., “future generation”).	Negative
Arming the Future	• There is a moral obligation to consider all options for future generations • Available CETs give future generations the ability to control the climate • Future generations should have the freedom to choose whether to use CETs	Compared to “Moral Hazard” and “Risk Transfer to the Future” positive future perspective on the research of CETs.	Positive
Technological Fix	• Technological fixes are attractive when citizens fail to make necessary behavioral changes • They are often simpler, faster, and require less effort than extensive social transformations • However, such solutions tinker with symptoms instead of resolving the causes, because it would permit continuing high levels of consumption, waste, and greenhouse gas emissions	Remark: ambivalent argument, depending if a respondent perceives a “Technological Fix” as something positive or negative (last bullet point).	Ambivalent

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=5272391

Table 1 (continued).

Maintaining the Status Quo	• CETs are a "pseudo-solution" that maintains the status quo and benefits industrial sectors and business branches that are the most reactionary in terms of climate policy • If CETs are controlled by big business, it may even reinforce the status quo • There is suspicion around the motivations, benefits, and secrecy of industries developing CETs	Respondents need to highlight in any form the “status quo”, which is perceived negatively.	Negative
Unstoppable Deployment if researched	• CETs research may generate internal momentum for deployment, even if unnecessary or not desirable • capital-intensive CETs would only be recouped over a long period of time • more investment in CETs research makes it harder to prevent future deployment	Compared to “Maintaining the Status Quo” this ethical argument emphasizes a *path dependency*, which are past decisions, which influence the choices and development of a system, often leading it down a specific trajectory, even when more efficient or rational alternatives may exist.	Negative
Emergency Case	• In case of a climate emergency (e.g., when climate sensitivity is high), CETs could stabilize temperatures • CETs could serve as a back-up plan or insurance against rapid, intense climate impacts • CETs could avert the worst effects of catastrophic climate events	Remark: ambivalent argument, depending if a respondent perceives SAI as a suitable technological fix in case of an emergency. Compared to the “Lesser-evil” this argument is more general and time-pressure is more decisive.	Ambivalent
Lesser-evil	• In a hypothetical scenario there may be a future situation where the deployment of CETs are necessary to prevent catastrophic climate change • In such a scenario the worst impacts of not deploying CETs may be worse than the risks associated with deploying it • CETs would be used as a last resort to avoid the worst impacts of climate change	In contrast to the “Buying Time” argument this argument emphasizes a negative hypothetical scenario and needs to include a comparison. The “Lesser-evil” argument should be often coupled with the “Emergency Case” argument.	Ambivalent

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=5272391

Table 1 (continued).

Buying Time	• CETs could be used as a temporary stopgap to buy time, e.g., for extending climate tipping points • CETs aims to bridge the gap until global mitigation policies become effective • CETs should only be time-limited until its goal is reached and should not lead to decreasing mitigation efforts	Compared to the “Lesser-evil” argument this argument is more general and highlights that CE should only buy time (paralleled by mitigation efforts) and / or should be limited.	Positive
Side-effects not predictable	• Uncertainties in CETs deployment cannot be substantially reduced through research • Deployment of these technologies is considered morally wrong due to these uncertainties • CETs may potentially worsen climate change instead of mitigating it (or increases human health risks)	Respondents could highlight here all kinds of possible side-effects, like increasing lung cancer, because of SAI, but not an unfair distribution of effects (“Unfair distribution of effects and power”).	Negative
Unfair distribution of effects and power	• CETs may disproportionately affect various communities and regions • This can result in unjust distributions of regional climate offsets, costs, and negative side-effects • Areas that have contributed least to climate change may bear most of these technologies' impacts	This argument highlights unfair distribution of effects and not unseen side-effects in general (“Side-effects unseen / not predictable”).	Negative
Hubris Argument	• Not engage in CETs, because the scope of the endeavor is beyond human understanding (virtue perspective) • CETs lack guaranteed effectiveness and full predictability of side effects (consequentialist perspective) • It demonstrates arrogance and self-deceit resulting from an unjustified confidence in knowledge and power beyond what is reasonable for humans	The argument can highlight the hubris for a single human (virtue) or the principle unpredictability of CETs side-effects (consequentialist) and leads to the conclusion not to engage in CETs (different to “Side-effects unseen / not predictable”).	Negative
Betrayal of Divine Creation	• Using CETs is a betrayal of Earth's purpose as given by a higher power (e.g., God). • CETs could signify a move toward "ending nature" and eliminating the world's inherent "wildness" (e.g., pure nature)	The argument highlights compared to the “Hubris Argument” a betrayal of a higher power / entity like God or the purity of nature.	Negative

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=5272391

Table 1 (continued).

Informed Consent	• CETs research and deployment require broad and well-informed consent • Consent should involve representatives of all potentially affected parties (just procedure) • All citizens have a legitimate stake in controlling the "global thermostat"	Remark: ambivalent argument, depending if a respondent perceives that an informed consent of all affected parties is possible. Respondents do not need to emphasize the second or last bullet point.	Ambivalent
Do it Alone	• A determined group of nations can deploy CETs, which benefit the entire world • Long-term cooperation or agreement from all nations may not be necessary	This argument emphasizes that the unilateral use of CETs is positive and not negative ("Risk of Unilateral Use").	Positive
Risk of Unilateral Use	• Research and development of CETs, especially SAI, may result in unilateral deployment with catastrophic consequences • Unilateral climate engineering can lead to political destabilization or be used for hostile purposes • CETs could even independently pursued by wealthy individuals or corporations	Compared to the "Do it Alone" argument this argument is negative and not explicitly highlighting the dual use of the technology as a potential weapon or strategic advantage ("Dual Use").	Negative
Dual Use	• CETs have the potential to modify the weather and therefore could be used as potential weapons • Nations could seek strategic advantage through climate modification methods	Compared to "Risk of Unilateral Use" the argument emphasizes that SAI could be used as a potential weapon or strategic advantage.	Negative
Risk of Governance	• Legal mechanisms for managing CETs, particularly SAI, pose a major challenge • A globally legitimate CETs regime would demand substantial geopolitical stability • SAI technology would need to be safeguarded against involuntary termination (e.g., by terrorist attacks)	The argument is quite broad, highlighting legal issues, geopolitical stability, or possible attacks on the SAI technology. However, when the issue of the long time frame is emphasized the "Long-Term Control" argument should be used.	Negative

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=5272391

Table 1 (continued).

Long-Term Control	• Social systems and institutions may struggle to control CETs over long time scales • Effective management is needed until greenhouse gas emissions are sufficiently reduced and SAI can be withdrawn	Compared to “Risk of Governance” the argument emphasizes the problem of long term control over time. Respondents do not need to emphasize the second bullet point.	Negative
Termination Problem (also “Not Addressing Root Problem”)	• In the absence of effective emissions reduction efforts, greenhouse gasses will continue to accumulate even if temperatures are artificially cooled through SAI • Therefor abrupt termination of SAI may result in rapid, catastrophic climate change, because of large concentration of atmospheric CO2	Remark: a potential termination problem exists only if insufficient mitigation efforts have been made. Therefore, SAI only treats symptoms (rising temperatures), but not causal problems (rising CO2 concentration).	Negative
Greater Good[a]	• If CETs are doing more good than harm then CETs should be deployed. • There could be a "moral obligation" or it could be in general "moral right" to use CETs (deontological perspective). • The technology is for the "greater good" or "maximizes benefits" for society (consequentialist perspective).	Compared to the “Lesser-evil” this argument is rather positive and no negative harms / side-effects are mentioned. There is no comparison (no hypothetical scenario). The argument can highlight the general obligation (deontological) or the positive consequences (consequentialist) using this technology.	Positive

Note. [a]This ethical argument was added after the first coding process (see below). Single ethical arguments like the last two arguments are only specific for SAI and not to “negative emissions” technologies in general.

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=5272391

Importantly, the evaluation of each single CET should consider its placement within a comprehensive climate portfolio, taking into account both the planned scale of its deployment and other climate options such as mitigation or adaptation measures (cf., Aldy et al., 2021; Gardiner, 2010; Neuber, 2018; Ott & Neuber, 2020). To inform such a comprehensive climate portfolio, we propose a study design and appropriate statistical procedures for online-studies in the following sections in order to empirically identify and investigate ethical concerns regarding SAI.

2. Study Design and Statistical Procedures

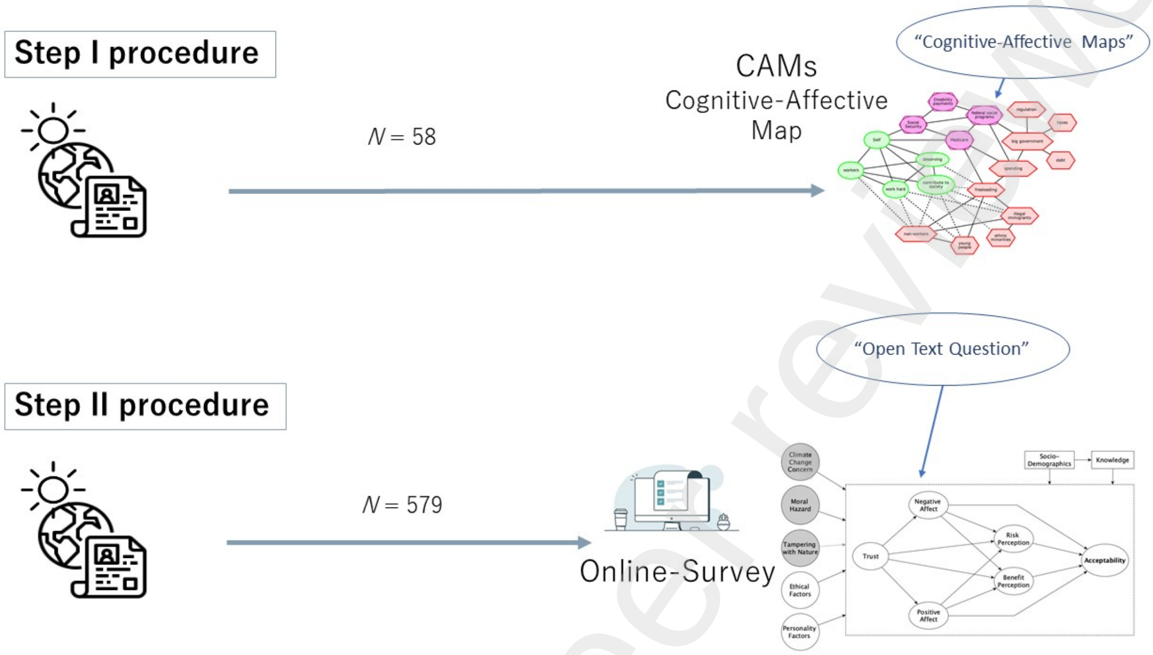
In the current paper, we re-analyse data of a previous study (Fenn et al., 2023), focusing on the identification of ethical concerns of laypersons regarding SAI based on two different types of data (see blue circles within Fig. 1). Participants were informed about the SAI technology by a pre-tested scenario text, describing possible benefits and risks (see https://osf.io/87w6g). The study was composed of two central steps: (a) CAMs were collected at the first measurement time point with a sample size of 58 participants. (b) At the second time point, a large-scale survey with a final sample size of 579 participants was conducted (for details, see Fenn et al., 2023).

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=5272391

Figure 1

Representation of the study design adapted from Fenn, et al. (2023), page 7.



Note. Within the circles, the two types of collected data analyzed in this article ("Cognitive-Affective Maps", "Open Text Question") are highlighted.

This complex study design allows for a multi-method approach to combine heterogeneous sources of data to inform the overall research question (cf., Johnson & Onwuegbuzie, 2004; Steegen et al., 2016). A variety of analytical methods were employed to process and interpret the data, with each method tailored to the specific data type, which are explained in more detail in the respective results sections.

2.1. Scenario-text approach

A balanced and pre-tested scenario text describing the operational principles and different advantages and disadvantages of the SAI technology was created (see Fenn et al., 2023). We considered it necessary to inform the participants about the SAI technology, because multiple articles have reported a relatively low knowledge regarding climate

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=5272391

engineering in general (e.g., Burns et al., 2016; Carlisle et al., 2020; Cummings et al., 2017; Merk et al., 2015). Importantly, we described SAI in the scenario text as imitating nature (e.g., by comparing the effect of SAI to that of volcanoes) to make the scenario text more easily understandable. However, this could have also influenced the perceived naturalness of the technology and could have artificially increased the acceptability (e.g., Corner & Pidgeon, 2015; Thomas et al., 2018). Such an effect is closely linked to the "Natural-is-better" heuristic, whereby nature mostly evokes positive emotions (Siegrist & Árvai, 2020; Siegrist & Hartmann, 2020).

2.2. Cognitive-Affective Maps

CAMs were collected in the Step I procedure (compare Fig. 1) in the study by Fenn et al. (2023). CAMs are a research method encompassing both qualitative and quantitative data-dimensions and can be viewed as a variant of mind maps (Reuter et al., 2022; Thagard, 2010). Participants used our recently developed tools (Fenn et al., under review)[3] to draw their CAM online. An exemplary CAM from the data set is shown in Fig. 3. A CAM consists of concepts and connections, freely drawn by participants to represent their associations. Each concept is assigned an affective connotation by participants on a scale ranging from [-3 to 3], indicating whether the concept evokes positive (green), negative (red), neutral (yellow), or ambivalent (purple) emotions. This visualization provides insights into the emotional valence associated with each concept as perceived by participants. Furthermore, it is possible to write comments to the drawn concepts to further elaborate the drawn concepts. Furthermore, CAMs permit to specify the strength and directionality of connections between these concepts. As such, CAMs can be described as a weighted directional network, which

[3] The interested reader is invited to try out our CAM tools online, as a detailed description of the tools is beyond the scope of this article, see: https://drawyourminds.de/

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=5272391

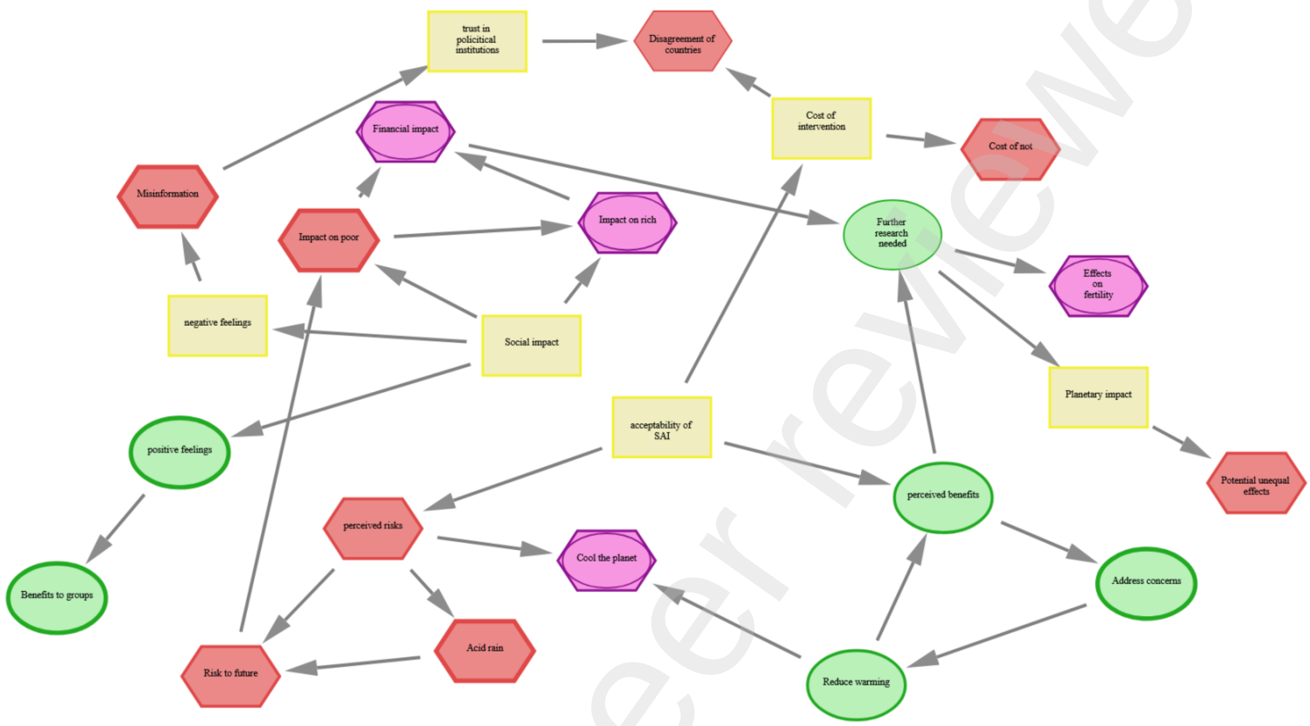
can be analyzed by procedures of network analysis (Bianconi, 2018; Fenn et al., under review; Newman, 2018), whereas the semantic content (written texts, comments) can be analyzed by means of Qualitative Content Analysis (Kuckartz & Rädiker, 2022; Mayring, 2022) and LLMs (Hussain et al., 2024; Tunstall et al., 2022).

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=5272391

Figure 3

Exemplary CAM with an average valence of -0.16 drawn by a participant.



Note. In this CAM, different concepts already indicate ethical arguments, which can be found in Tab. 1. For example, “Disagreement of countries” corresponds to the ethical argument of “Risk of Governance”.

2.3. Open-Text

In the Step II procedure in the study by Fenn et al. (2023), participants first read the scenario text and then immediately answered the following question: “When, in your opinion, is the described ‘Stratospheric Aerosol Injection’ technology morally right?”. Additionally, participants were provided a general definition of morality (based on Jacobs, 2002; Pieper, 2017). Participants were forced to take at least one minute to answer this open text question. These open text answers were analyzed according to the procedure of qualitative content analysis. To support the raters to code the text material applying qualitative content analysis,

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=5272391

a YouTube video with coding instructions was created[4]. Qualitative content analysis is a systematic method of analyzing text data that involves coding and categorizing the content to derive themes and patterns. It emphasizes a structured approach that includes several stages, such as preparation, forming categories, and coding the material (Kuckartz & Rädiker, 2022; Mayring, 2022). To code the text material, coding guidelines which contain all the theoretically derived ethical arguments (see Tab. 1) were developed. Then, a three step coding procedure followed (see "Results" section for details).

3. Results

The following statistical analyses were analyzed using R (R Core Team, 2020), Mplus (Muthén & Muthen, 2017) and Python (Van Rossum & Drake, 2009). The analysis scripts, which have been written in the form of text annotated reproducible scripts by using the "rmarkdown" package in R (Xie et al., 2018) and the Jupyter notebook for Python (Kluyver et al., 2016), are publicly accessible via the Open Science Framework (OSF) at https://doi.org/10.17605/OSF.IO/ANXG6.

The subsequent sections employ a deductive-driven qualitative content analysis to summarize insights from the CAMs and open-text responses. In contrast, the final results section adopts a predominantly inductive, data-driven approach, leveraging LLMs to synthesize the most prevalent ethical arguments and underlying patterns of reasoning. Key findings are summarized in Table 2 presented in the "Discussion" section.

3.1. Cognitive-Affective Maps

58 participants (mean age 38, *SD* = 10.40, 47% female) drew the CAMs online by using recently developed tools (Fenn et al., 2024). We pre-defined the concept "acceptability

[4] see YouTube Video: https://www.youtube.com/watch?v=725flcytGJw

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=5272391

of SAI" in the center of the CAM. On the top, five additional concepts ("positive feelings", "negative feelings", "trust in political institutions", "perceived risks", and "perceived benefits") were presented. Participants were able to move or delete the five additional pre-defined concepts and were technically required to draw at least 24 concepts in total (see in detail Fenn et al., 2023). The dataset consists of 58 CAMs, where participants drew an average of 25.4 concepts each, with 34% positive, 46% negative, 12% neutral, and 8% ambivalent. On average, 44.21 connectors were drawn per CAM. The mean valence across all CAMs was -0.33, and at least one predefined concept was removed in 14% of the CAMs. Detailed descriptive statistics are provided in Appendix A.

3.1.1. Data preparation

While summarizing the CAM data for this article, we focused on ethical arguments within the individually drawn CAMs. Thereby, each drawn concept can be assigned to one of the 21 different ethical arguments (see Tab. 1). The CAM data were summarized by our developed Data Analysis Tool, and we iteratively summarized the semantic content of the CAMs (Fenn et al., 2023, under review). The initial 1,473 concepts (consisting of 1,063 unique concepts) were condensed to a final set of 41 concepts.

3.1.2. Data analysis

For this analysis, we focus only on ethical arguments (see Tab. 1) within the CAMs. 19 of the 21 ethical arguments were mentioned at least by one participant – only two ethical arguments "Emergency Case" and "Unstoppable Deployment if researched" were not mentioned. Overall 724 of the total 1,473 drawn concepts could be assigned to ethical arguments, whereby the argument "Side-effects not predictable" was mentioned most frequently (264 times), followed by the positive ethical argument "Greater Good" (143). The ethical argument of "Side-effects not predictable" encompasses all kinds of negative

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=5272391

evaluations, e.g., SAI could cause “acid rain”, increase “health risks”, and “side-effects [are] not fully resolved”. In contrast, “Greater Good” highlights that SAI could have “environmental benefits” or could even “save the world”. Participants also frequently highlighted the “Risk of Governance” (e.g., unclear who is “accountable” or risks of “political instability”), which is linked to the negatively perceived argument of “Informed Consent” highlighting the lack of “political consensus” and the impossibility “to get everyone to agree on [SAI]”. The argument of “Long-Term Control” is also negatively perceived, conveying that it “could be hard to get all countries to commit for a long period of time”. However, there is a mixed perception regarding SAI as a “Technological Fix”. Some participants viewed it as “a solution where we don’t have to make changes to our everyday life,” while more frequently, participants highlighted concerns that SAI could foster a “Moral Hazard”. A table with the frequencies of all the mentioned ethical arguments and a few examples from participants can be found in Appendix B, a complete wordlist can be found on OSF (https://osf.io/rhxfn).

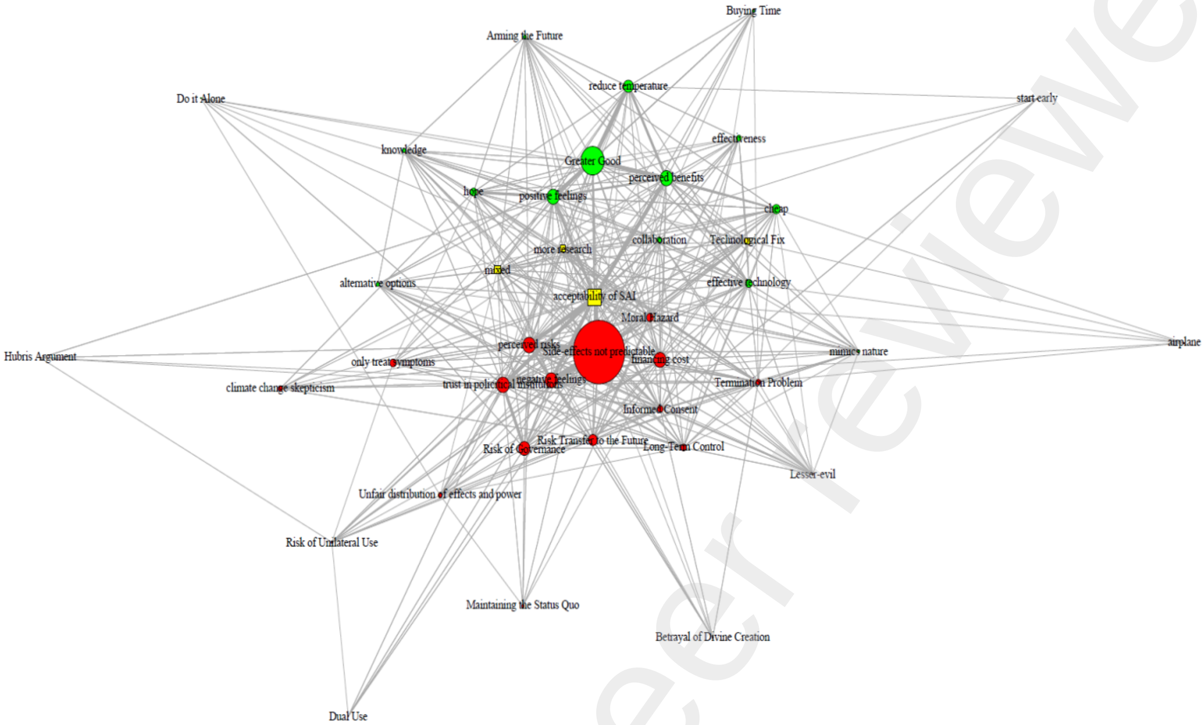
All these ethical arguments are interrelated, which can be seen in Fig. 4, where we show an aggregated network based on all 58 CAMs. The relative size of the concepts and the thickness of the connections indicate the frequency of the drawn concepts and the frequency of the pairwise connections, respectively. On average, participants mentioned 12.48 (*SD* = 3.40) ethical arguments in their CAMs.

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=5272391

Figure 4

Aggregated CAM consisting of N = 58 CAMs.



Note. A zoomable PDF file can be found on OSF (https://osf.io/xr62c). The color is indicative of the average valence of the concepts, whereby yellow represents neutral, green positive, and red negative concepts. If the average valence of a concept is within [-0.5, 0.5] a concept was drawn as neutral. Remark: Figure was adjusted from Fig. 3 in Fenn et al. (2023) on page 11.

3.1.3. Discussion of CAM results

Most frequently, participants highlighted negative ethical arguments, especially all kinds of possible side-effects. As visible in Fig. 4, these ethical arguments are strongly interrelated, whereby participants emphasize the importance of governance related ethical arguments ("Risk of Governance", "Informed Consent", "Long-Term Control", "Termination Problem"). Such concerns could inform possible issues of governing solar geoengineering (e.g., Flegal et al., 2019; MacMartin et al., 2019). As can be seen in the upper part of Fig. 4, participants drew multiple positively assessed concepts regarding SAI (e.g., "Greater Good").

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=5272391

These concepts are connected to other summarized concepts, like SAI being capable of reducing the temperature or being relatively cheap. About 9% of participants mentioned "Betrayal of Divine Creation" as strongly negative, arguing that SAI is not acceptable because SAI is like "playing God" or violates the purity of nature. Such an argument could be a strong moral heuristic when making (ethical) decisions (cf., Schwartz, 2016). Also, 16% of the participants drew positive concepts that SAI is mimicking nature, e.g., that SAI "would have similar effects on the atmosphere as volcanoes" (see Appendix B). Such a finding emphasizes the importance of how CETs are framed in general (see section 2.1.).

3.2. Open Text

In total, we had a final sample size of 579 participants (M = 40 years, SD = 13.26, 47% female). Three participants provided no answer and two participants indicated that they did not know how to answer the open text question regarding the morality of SAI. Removing participants with no answers we have 576 answers varying in length from 1 to 171 words (mean number of words: 36.26, SD = 21.96). By applying the Python module VADER (for Valence Aware Dictionary for sEntiment Reasoning; Hutto & Gilbert, 2014), we computed sentiment scores which refer to the emotional tone expressed in the text answers. In total, 273 answers were negative, 246 positive and 57 neutral. Descriptively, neutral arguments seemed less elaborated and had on average only 17.79 words (SD = 11.35). The following section outlines the three-step procedure used to categorize ethical arguments within the text responses.

3.2.1. Data preparation

The summary of the open text data was based on qualitative content analysis (Mayring, 2022) using the open access QCAmap application (Fenzl & Mayring, 2017). To summarize the data, we followed the strict step model of the procedure of "deductive

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=5272391

category assignment"[5]. Based on existing theories, a category system was defined (see Tab. 1), followed by seven raters coding 5% of the text answers in a first step. This procedure led to minor adjustments of the category system. Importantly, a new ethical argument ("Greater Good") was added to code text answers emphasizing that the technology is doing more good than harm without mentioning negative side-effects (for details see: https://osf.io/evzwm). Finally, in the last step, the complete text data was coded by seven raters, whereby six of these raters also participated in the first coding process.

Multiple quality criteria were applied to check the quality of the summarizing process of the text data (motivated by Kuckartz & Rädiker, 2022; Mayring, 2022). The content validity (Moosbrugger & Kelava, 2020) of the category system was reflected within team discussions involving two ethics experts to ensure that the ethical arguments (categories) were constructed in such a way that they capture central reasoning discussed in the philosophy of ethics. For the first coding process, we tested for the inter-rater reliability, computing Fleiss' kappa (Fleiss et al., 2013) to check for discrepancies between the seven raters for every single open text answer coded. On average, the reliability was substantial with $\varkappa = .75$ ($p < .01$). To improve the coding of the complete text data (final step), we followed the procedure of "subjective assessment" (Guest et al., 2012), whereby discrepancies were discussed in a group discussion with all the raters until consensus was reached. The category system was adjusted respectively.

3.2.2. Data analysis

In the following, only 553 of 576 text answers are considered, because in 15 (2.6%) text answers no ethical argument was assigned and in 11 (1.9%) no consensus was found regarding the coding of the respecting text answer. Motivated by Pokorny et al. (2018), we

[5] Accessible in the QCAmap application, see: https://www.qcamap.org/ui/assets/tutorials/en/Steps_Rules_Deductive.pdf; last accessed on January 9, 2025

Preprint not peer reviewed

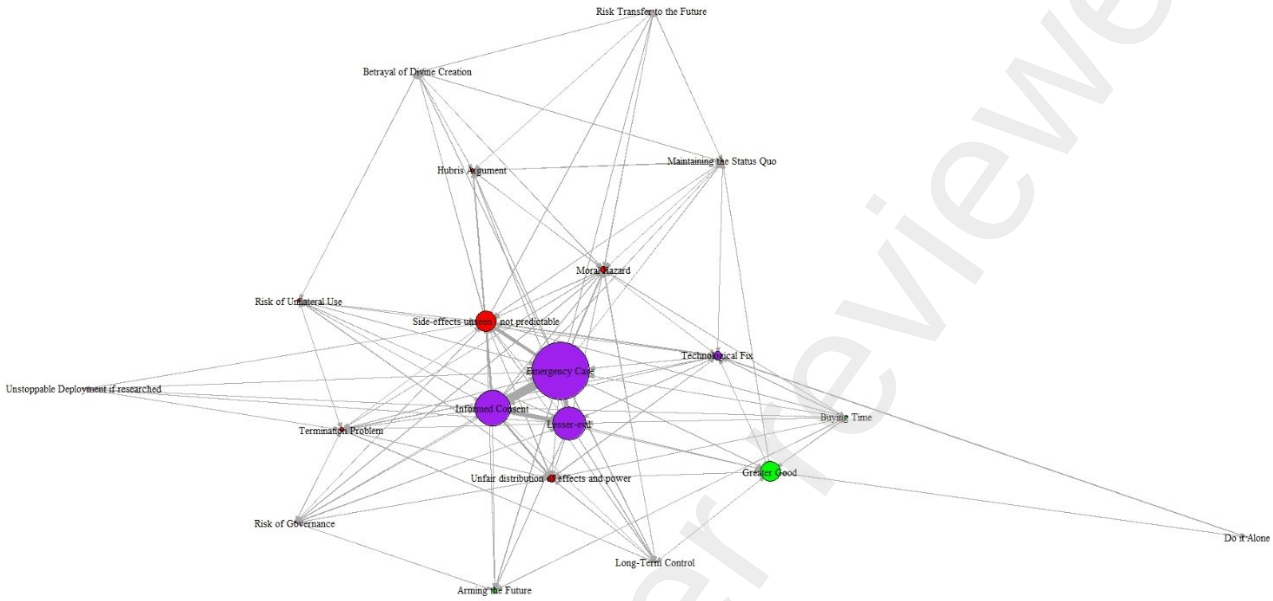
This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=5272391

visualized the coding of the 553 text answers in the form of a network in Fig. 5. The three most frequently mentioned ethical arguments were “Emergency Case” (assigned 202 times), followed by “Informed Consent” (126) and “Lesser-Evil” (119). These three ambivalent ethical arguments (see Tab. 1) were often mentioned within the same text answers, linking them argumentatively. For example, a participant linked the ethical arguments of “Emergency Case” and “Lesser-Evil” by stating if “climate change has reached a point where human life is affected to such a degree that large numbers of people are facing hardships [..] [SAI] would be a last resort” (quote of one participant). Using this technology in “an imminent catastrophic climate emergency” would only be morally right if this action is “agreed on by all the major or world leading countries of the world” (quote by another participant) highlights a connection between “Lesser-Evil” and “Informed Consent”. Thereby, open text answers varied regarding who needs to agree (e.g., global, all affected, or leading countries; see word clouds of given answers: https://osf.io/5ztcj). Interestingly, participants also frequently mentioned the “Greater Good” of the technology, expressing a generally positive attitude towards the technology, e.g., “it will benefit society, and help in general” (quote of another participant). In addition, the argument of the general unpredictability of the technology (ethical argument of “Side-effects not predictable”) was mentioned as often as the "Greater Good" argument. The ethical argument of “Dual Use” has not been mentioned by a single participant and some ethical arguments were mentioned less than ten times (for details see table in Appendix C).

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=5272391

Figure 5

Covariation of ethical arguments within open text.



Note. A zoomable PDF file can be found on OSF (https://osf.io/mxt89). The color coding represents the "General Evaluation" column in Tab. 1 with green = positive, red = negative, yellow = neutral, and purple = ambivalent. The frequency of drawn concepts and the number of pairwise connections is proportional to the size of the concept and the thickness of the connections, respectively.

3.2.3. Discussion of Open Text results

Participants mentioned ambivalent ethical arguments most frequently in their open text answers, frequently highlighting that SAI would be the "Lesser-Evil" if there would be an "Emergency Case". At the same time, it was also emphasized that the informed consent of all, or at least of the majority of countries, would be necessary. Unlike the CAM data, participants highlighted possible negative-side effects and general benefits of the SAI technology ("Greater Good") to an equal extent. Also, governance-related issues were hardly mentioned. Thus, ambivalent ethical arguments strongly dominated the text answers.

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=5272391

Importantly, multiple participants already emphasized that the ethical arguments of "Emergency Case" and "Lesser-Evil" are strongly related (e.g., Gardiner, 2010b; Ott & Neuber, 2020), whereby the lesser evil argument has the potential to become a self-fulfilling prophecy, as preparing for a horrific scenario may inadvertently lead to its occurrence.

3.3. Testing Large Language Models for Summarizing Ethical Arguments

To analyze coding assignments derived from the CAM alongside open-text data, we utilized a LLM, specifically "Llama-3.1-70B-Instruct" (Dubey et al., 2024) , in two distinct applications: First, the model was tasked with synthesizing marked text passages - segments of text identified as relevant during the coding process - to generate comprehensive summaries of the corresponding ethical arguments. This initial application facilitated the distillation of extensive textual datasets into concise, interpretable summaries tailored to the ethical coding guidelines under examination. Subsequently, the LLM was applied again to identify commonalities and differences by composing a synthesized paragraph summarizing the shared themes and discrepancies between our formal definition of the ethical arguments (see Table 1) and laypersons' associations with the respective ethical argument. Both times, we set up well-structured prompts by including contextual data, explanations of the data structure, and including the respective definition of the ethical argument into the prompts, adhering to established best practices in the literature (Dai et al., 2023; Liu et al., 2023; White et al., 2023). A non-technical introduction to LLMs is provided in Appendix D.

3.3.1. LLM-Generated Summaries of Ethical Arguments

Table 2 presents the outcomes of the initial prompting of the LLM, offering a structured and concise synthesis of two exemplary ethical arguments: *Moral Hazard* and *Technological Fix* (for all other ethical arguments see table on OSF: https://osf.io/jx4hc**)**. For each of the two

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=5272391

datasets – Open Text and CAM data – the LLM-generated summaries encapsulate the most salient themes and patterns identified within the diverse associations articulated by laypersons in response to these ethical arguments.

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=5272391

Table 2

Laypersons' Associations with Ethical Arguments on SAI, Synthesized Using LLM

Ethical Argument	Summary Open Text	Summary CAM
Moral Hazard (also: "Undermining Better Options")	• CE might reduce enthusiasm for mitigation efforts, as it may be viewed as a last resort for continued pollution. • Focusing on CE research may divert resources from essential emission abatement and adaptation strategies. • CE does not address climate change's root cause (CO_2 emissions) and may lead to complacency. • Tackling excessive CO_2 emissions through reduced energy use, green energy, and lifestyle changes is crucial. • Prioritizing CE over mitigation efforts is morally problematic, as it doesn't resolve pollution or high CO_2 levels.	• CE could lead to complacency, as people may view it as a solution and reduce efforts to address climate change. • It may be seen as a distraction from real issues, diverting resources away from long-term sustainable change. • Governments may choose CE for its simplicity, reducing investment in sustainable energy and climate action. • CE doesn't encourage innovation and may lead to political lethargy, preventing necessary action against climate change.
Technological Fix	• Technological fixes like SAI are justified only after all other options have been explored, as they treat symptoms, not the root cause. • Without addressing CO_2 emissions and behavioral changes, such fixes are counterproductive and not sustainable. • Reducing CO_2 emissions and promoting behavioral changes should be prioritized over temporary technological fixes. • Technological fixes can lead to complacency, reducing motivation for necessary systemic changes. • Such solutions should be last resorts, used only when efforts to reduce emissions fail.	• Technological fixes like SAI are seen as quick solutions but may delay necessary lifestyle and behavioral changes. • Adoption may be driven by political indecision or lack of will to implement more comprehensive climate solutions. • While urgent, technological fixes don't address the root cause of climate change and may offer only temporary relief. • Their low cost and simplicity may influence decisions, but they are not always the most effective long-term solution. • They may reduce personal responsibility, offering an easy way out instead of encouraging behavioral change.

Note. All other layperson's associations to the ethical arguments can be found on OSF, see: https://osf.io/jx4hc

3.3.2. Exploring Ethical Parallels: Formal Definitions vs. Lay Perspectives

The outcome of this LLM prompt is a synthesized and concise summary that systematically identifies and highlights the shared themes and significant discrepancies between laypersons' interpretations and our formal definitions of the ethical arguments (see table available on OSF: https://osf.io/sncjh). This comparative synthesis highlights notable differences between our theory-driven definitions of the ethical arguments (see Table 1) and laypersons'

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=5272391

associations with these arguments: These differences arise from contrasting emphases on moral, social, and technical dimensions. In our definitions we predominantly focused on structured conceptual frameworks, such as resource diversion (Moral Hazard), future dilemmas (Risk Transfer), and the theoretical last-resort nature of climate engineering (Lesser Evil). Conversely, laypersons emphasize immediate moral implications, societal trust, and tangible impacts. Lay perspectives especially highlight issues such as climate engineering's failure to address root causes, political inertia, global conflicts, and inter-nation mistrust, alongside the need for systemic change and accountability. Context-specific concerns - such as temperature thresholds (Emergency Case) and catastrophic consequences (Risk Transfer) - also feature more prominently in laypersons' interpretations. Overall laypersons' arguments are framed through lived experiences and practical concerns, often introducing notions of fairness, equity, and societal trust.

3.3.3. Discussion of LLM results

A comparison of the LLM-generated summaries in Section 3.3.1 reveals both similarities and differences in the emphases of CAM and open-text data. Both datasets consistently highlight the moral and ethical implications of CE sharing concerns, for example, about its potential to divert attention from addressing root causes and undermining sustainable climate solutions. However, CAM data more frequently emphasizes geopolitical and social dimensions, such as global conflicts and inter-nation mistrust, whereas open-text data focuses on individual responsibility and the moral necessity of addressing systemic issues like resource diversion and behavioral change. The greater emphasis on geopolitical dimensions in CAM data could be attributed to the predefined inclusion of the concept "trust in political institutions", which shaped the framing of responses in that dataset. As shown in Section 3.3.2., notable

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=5272391

discrepancies exist between laypersons' interpretations and our formal definitions of ethical arguments. These differences reflect contrasting emphases: while formal definitions focus on structured, theoretical constructs such as resource diversion and risk transfer, laypersons prioritize immediate moral implications, lived experiences, and tangible outcomes. These results underscore the unique contribution of laypersons' associations in expanding the discourse around ethical arguments.

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=5272391

4. Conclusion

In this article, we demonstrate the value of integrating two distinct data types – open text responses and CAMs – to explore laypersons' ethical concerns regarding the use of SAI. CAMs offer a structured visualization of ethical concerns, identifying a broad spectrum of issues ranging from "trust in political institutions" to "mimicking nature". Thereby participants structure ethical arguments in the process of drawing CAM, which is related to the theoretical concept of ethical coherence (Thagard, 1998, 2000). In contrast, open-text data revealed mainly ambivalent arguments (e.g., "Emergency Case," "Lesser-Evil," "Informed Consent"). Open text and CAMs could enable future researchers to identify central ethical arguments or even *master-narratives* regarding (such) emerging technologies, such as the notion that deploying these technologies is like "Opening Pandora's box" (Davies & Macnaghten, 2010; Macnaghten et al., 2019). It could be the case that CAMs foster deliberative thinking and enable participants to structure complex ethical arguments in the form of complex interconnected maps (see Vink et al., 2016). Methodological differences are summarized alongside our key findings in Table 3.

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=5272391

Table 3

Overview of the type of data, the main outcomes and reflection of the results of the two types of data

	Cognitive-Affective Maps	Open Text
Type of Data	qualitative and quantitative	qualitative
Main Outcomes	* broad range of ethical arguments identified, including governance related arguments * ethical arguments are linked to other predefined concepts (e.g., “trust in political institutions”) * arguments like “feeling that SAI mimics nature” or “brings hope” could influence the ethical argumentation	* mainly ambivalent ethical arguments identified * three ambivalent ethical arguments (“Emergency Case”, “Lesser-Evil”, “Informed Consent”) are argumentatively interlinked in the text answers
Reflection	* depending on the pre-defined concepts, participants probably highlight different ethical arguments * participants were required to draw 24 concepts, potentially leading to the high number of possible negative side-effects that were mentioned	* by answering a general question regarding the morality of SAI, participants might be inclined to think about the argument that SAI could be used in case of an emergency * ethical arguments, like governance issues, do not seem to be mentally present

Note. The “Type of Data” can be “qualitative” (e.g., answer to open-ended survey questions resulting in text) or “quantitative” (e.g., drawing a network resulting in specific network parameters).

Referring to the detailed review by Reynolds & Horton (2020) the findings outlined in this article yield insights for the analytical problems of the Earth System Governance framework (Biermann et al., 2010; Burch et al., 2019): Laypersons in the CAM data highlighted governance related ethical arguments and emphasized central problems like inter-nation mistrust, unclear accountability, lack of consensus or potential political instability highlighting problems of a potential future governance architecture. Further the open-text data revealed concerns about equitable participation (“informed consent of all countries”) and thereby pointing to the moral legitimacy of decision-making, emphasizing concerns of potential power asymmetries. The frequent “Lesser-Evil” and “Emergency Case” arguments underscore distributive and intergenerational justice challenges, while the LLM-mediated

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=5272391

synthesis emphasizes how lay perspectives foreground practical burdens and potential benefits for vulnerable populations. Lastly laypersons imagined future scenarios, whether hopeful ("Greater Good") or alarming ("Betrayal of Divine Creation"), illustrating the power of narrative imaginations.

4.1. Future research and limitations

In our opinion, the ethical concerns of emerging CETs should be assessed continuously, because if an ethical assessment would wait until sufficient information regarding the (side-)effects of a CET becomes available, a technology would be deeply ingrained in society and the potential for making revisions would be strongly limited (called the "Collingridge dilemma", Collingridge, 1980; Möller & Grießhammer, 2022). This justifies an early ethical assessment of such technologies, even if knowledge of laypersons regarding CETs is low (Grunwald & Hillerbrand, 2021; Palm & Hansson, 2006). Such a perspective emphasizes the need to conduct future studies, e.g., ideally as a "tracker technology assessment" (see Böschen et al., 2021; Lucivero, 2016) to inform the empirical ethical assessment of CETs during different development stages of these technologies. For example, "Moral Hazard" is of particular importance at an early research stage while the "Termination Problem" is particularly important when such a technology would be finally implemented (see Preston, 2013). Future research should systematically examine whether increasing participants' knowledge about CETs influences their ethical concerns and their envisioning of potential futures involving such technologies. Here one might conjecture, for example, information-choice questionnaires (e.g., Gregory et al., 2016; ter Mors et al., 2013). In addition, a future study could provide a more straightforward comparison of CAMs and text data. In the present study, such a comparison was limited due to fundamental differences in methodological design - particularly the pre-defined concepts in the CAM approach, which

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=5272391

may trigger different cognitive associations and thereby shape the ethical concerns participants express (see Lichtenstein & Slovic, 2006). If CAMs provide similar information as text data, we would recommend applying CAMs with different sets of pre-defined concepts because such data can be semi-automatically summarized.

To inform a comprehensive climate portfolio, the ethical concerns of all important CETs should be assessed by multiple stakeholder groups (cf., Aldy et al., 2021; Gardiner, 2010; Neuber, 2018; Ott & Neuber, 2020). We therefore encourage future researchers to adopt the methodology proposed in this study, along with the accompanying online resources, to advance empirically informed ethics of CETs. Incorporating laypersons' perspectives can enhance the inclusivity and societal relevance of discourse on these technologies, thereby supporting climate policy and anticipatory governance.

Finally we want to stress that research in this domain is particularly needed, as participants in our study referred to the potential deployment of SAI in the context of a future climate emergency, framing it as a "lesser evil" in a hypothetical but severe crisis scenario. This justification for the ultimate use of such a risky technology (e.g., Sovacool et al., 2022, 2023) underscores the critical importance of preemptively avoiding such "emergency situations" so that there is still room for ethical discussions to govern such technologies before potentially irreversible measures become necessary (cf., Gardiner, 2011; Ott, 2011).

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=5272391

References

Aldy, J. E., Felgenhauer, T., Pizer, W. A., Tavoni, M., Belaia, M., Borsuk, M. E., Ghosh, A., Heutel, G., Heyen, D., Horton, J., Keith, D., Merk, C., Moreno-Cruz, J., Reynolds, J. L., Ricke, K., Rickels, W., Shayegh, S., Smith, W., Tilmes, S., … Wiener, J. B. (2021). Social science research to inform solar geoengineering. *Science*, *374*(6569), 815–818. https://doi.org/10.1126/science.abj6517

Anderson, K., & Peters, G. (2016). The trouble with negative emissions. *Science*, *354*(6309), 182–183. https://doi.org/10.1126/science.aah4567

Barrett, S. (2008). The Incredible Economics of Geoengineering. *Environmental and Resource Economics*, *39*(1), 45–54. https://doi.org/10.1007/s10640-007-9174-8

Betz, G., & Cacean, S. (2012). *Ethical Aspects of Climate Engineering*. KIT Scientific Publishing.

Bianconi, G. (2018). *Multilayer Networks: Structure and Function*. Oxford University Press.

Biermann, F., Betsill, M. M., Gupta, J., Kanie, N., Lebel, L., Liverman, D., Schroeder, H., Siebenhüner, B., & Zondervan, R. (2010). Earth system governance: A research framework. *International Environmental Agreements: Politics, Law and Economics*, *10*(4), 277–298. https://doi.org/10.1007/s10784-010-9137-3

Biermann, F., Oomen, J., Gupta, A., Ali, S. H., Conca, K., Hajer, M. A., Kashwan, P., Kotzé, L. J., Leach, M., Messner, D., Okereke, C., Persson, Å., Potočnik, J., Schlosberg, D., Scobie, M., & VanDeveer, S. D. (2022). Solar geoengineering: The case for an international non-use agreement. *WIREs Climate Change*, *13*(3), e754. https://doi.org/10.1002/wcc.754

Bindé, J. (2000). Toward an Ethics of the Future. *Public Culture*, *12*(1), 51–72. https://doi.org/10.1215/08992363-12-1-51

Böschen, S., Grunwald, A., Krings, B.-J., & Rösch, C. (2021). *Technikfolgenabschätzung:*

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=5272391

Handbuch für Wissenschaft und Praxis. Nomos Verlag.

Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., Agarwal, S., Herbert-Voss, A., Krueger, G., Henighan, T., Child, R., Ramesh, A., Ziegler, D. M., Wu, J., Winter, C., … Amodei, D. (2020). *Language Models are Few-Shot Learners* (No. arXiv:2005.14165). arXiv. https://doi.org/10.48550/arXiv.2005.14165

Buckley, J. A., Thompson, P. B., & Whyte, K. P. (2017). Collingridge's dilemma and the early ethical assessment of emerging technology: The case of nanotechnology enabled biosensors. *Technology in Society*, *48*, 54–63. https://doi.org/10.1016/j.techsoc.2016.12.003

Burch, S., Gupta, A., Inoue, C. Y. A., Kalfagianni, A., Persson, Å., Gerlak, A. K., Ishii, A., Patterson, J., Pickering, J., Scobie, M., Van der Heijden, J., Vervoort, J., Adler, C., Bloomfield, M., Djalante, R., Dryzek, J., Galaz, V., Gordon, C., Harmon, R., … Zondervan, R. (2019). New directions in earth system governance research. *Earth System Governance*, *1*, 100006. https://doi.org/10.1016/j.esg.2019.100006

Burns, E. T., Flegal, J. A., Keith, D. W., Mahajan, A., Tingley, D., & Wagner, G. (2016). What do people think when they think about solar geoengineering? A review of empirical social science literature, and prospects for future research. *Earth's Future*, *4*(11), 536–542. https://doi.org/10.1002/2016EF000461

Caelen, O., & Blete, M.-A. (2023). *Developing Apps with Gpt-4 and Chatgpt: Build Intelligent Chatbots, Content Generators, and More* (1st ed.). O'Reilly Media.

Callies, D. E. (2019). The Slippery Slope Argument against Geoengineering Research. *Journal of Applied Philosophy*, *36*(4), 675–687. https://doi.org/10.1111/japp.12345

Carlisle, D. P., Feetham, P. M., Wright, M. J., & Teagle, D. A. H. (2020). The public remain uninformed and wary of climate engineering. *Climatic Change*, *160*(2), 303–322.

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=5272391

https://doi.org/10.1007/s10584-020-02706-5

Carr, W. A., & Yung, L. (2018). Perceptions of climate engineering in the South Pacific, Sub-Saharan Africa, and North American Arctic. *Climatic Change*, *147*(1), 119–132. https://doi.org/10.1007/s10584-018-2138-x

Caviezel, C., & Revermann, C. (2014). *Climate Engineering: Kann und soll man die Erderwärmung technisch eindämmen?* edition sigma. https://publikationen.bibliothek.kit.edu/140100230

Chiang, W.-L., Zheng, L., Sheng, Y., Angelopoulos, A. N., Li, T., Li, D., Zhang, H., Zhu, B., Jordan, M., Gonzalez, J. E., & Stoica, I. (2024). *Chatbot Arena: An Open Platform for Evaluating LLMs by Human Preference* (No. arXiv:2403.04132). arXiv. https://doi.org/10.48550/arXiv.2403.04132

Collingridge, D. (1980). *Social Control of Technology*. Milton Keynes: Open University Press.

Corner, A., Parkhill, K. A., & Pidgeon, N. (2011). *"Experiment Earth?" Reflections on a public dialogue on geoengineering: Reflections on a public dialogue on geoengineering* [Working Paper]. Cardiff University. https://eprints.whiterose.ac.uk/82861/

Corner, A., Parkhill, K., Pidgeon, N., & Vaughan, N. E. (2013). Messing with nature? Exploring public perceptions of geoengineering in the UK. *Global Environmental Change*, *23*(5), 938–947. https://doi.org/10.1016/j.gloenvcha.2013.06.002

Corner, A., & Pidgeon, N. (2014). Geoengineering, climate change scepticism and the 'moral hazard' argument: An experimental study of UK public perceptions. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *372*(2031), 20140063. https://doi.org/10.1098/rsta.2014.0063

Corner, A., & Pidgeon, N. (2015). Like artificial trees? The effect of framing by natural

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=5272391

analogy on public perceptions of geoengineering. *Climatic Change*, *130*(3), 425–438. https://doi.org/10.1007/s10584-014-1148-6

Cotton, M. (2014). *Ethics and Technology Assessment: A Participatory Approach*. Springer.

Crutzen, P. J. (2006). Albedo Enhancement by Stratospheric Sulfur Injections: A Contribution to Resolve a Policy Dilemma? *Climatic Change*, *77*(3–4), 211. https://doi.org/10.1007/s10584-006-9101-y

Cummings, C. L., Lin, S. H., & Trump, B. D. (2017). Public perceptions of climate geoengineering: A systematic review of the literature. *Climate Research*, *73*(3), 247–264. https://doi.org/10.3354/cr01475

Dai, S.-C., Xiong, A., & Ku, L.-W. (2023). *LLM-in-the-loop: Leveraging Large Language Model for Thematic Analysis* (No. arXiv:2310.15100). arXiv. https://doi.org/10.48550/arXiv.2310.15100

Davies, S. R., & Macnaghten, P. (2010). Narratives of Mastery and Resistance: Lay Ethics of Nanotechnology. *NanoEthics*, *4*(2), 141–151. https://doi.org/10.1007/s11569-010-0096-5

De Bruin, K. C., Dellink, R. B., & Tol, R. S. J. (2009). AD-DICE: An implementation of adaptation in the DICE model. *Climatic Change*, *95*(1–2), 63–81. https://doi.org/10.1007/s10584-008-9535-5

Debelak, R., Koch, T., Aßenmacher, M., & Stachl, C. (2024). *From Embeddings to Explainability: A Tutorial on Transformer-Based Text Analysis for Social and Behavioral Scientists*. OSF. https://doi.org/10.31234/osf.io/bc56a

Dowling, A. (2018). *Greenhouse Gas Removal* [Monograph]. Royal Society. https://royalsociety.org/topics-policy/projects/greenhouse-gas-removal/

Dubey, A., Jauhri, A., Pandey, A., Kadian, A., Al-Dahle, A., Letman, A., Mathur, A., Schelten, A., Yang, A., Fan, A., Goyal, A., Hartshorn, A., Yang, A., Mitra, A.,

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=5272391

Sravankumar, A., Korenev, A., Hinsvark, A., Rao, A., Zhang, A., … Zhao, Z. (2024). *The Llama 3 Herd of Models* (No. arXiv:2407.21783). arXiv. https://doi.org/10.48550/arXiv.2407.21783

Fenn, J., Gouret, F., Gorki, M., Reuter, L., Gros, W., Hüttner, P., & Kiesel, A. (under review). *Cognitive-Affective Maps extended logic: Proposing Tools to Collect and Analyze Attitudes and Belief Systems*.

Fenn, J., Helm, J. F., Höfele, P., Kulbe, L., Ernst, A., & Kiesel, A. (2023). Identifying key-psychological factors influencing the acceptance of yet emerging technologies–A multi-method-approach to inform climate policy. *PLOS Climate*, *2*(6), 1–25. https://doi.org/10.1371/journal.pclm.0000207

Fenzl, T., & Mayring, P. (2017). *QCAmap: Eine interaktive Webapplikation für Qualitative Inhaltsanalyse*. https://doi.org/10.23668/psycharchives.11259

Fiorino, D. J. (1990). Citizen Participation and Environmental Risk: A Survey of Institutional Mechanisms. *Science, Technology, & Human Values*, *15*(2), 226–243. https://doi.org/10.1177/016224399001500204

Flegal, J. A., Hubert, A.-M., Morrow, D. R., & Moreno-Cruz, J. B. (2019). Solar Geoengineering: Social Science, Legal, Ethical, and Economic Frameworks. *Annual Review of Environment and Resources*, *44*(1), 399–423. https://doi.org/10.1146/annurev-environ-102017-030032

Fleiss, J. L., Levin, B., & Paik, M. C. (2013). *Statistical Methods for Rates and Proportions*. John Wiley & Sons.

Frumhoff, P. C., & Stephens, J. C. (2018). Towards legitimacy of the solar geoengineering research enterprise. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *376*(2119), 20160459. https://doi.org/10.1098/rsta.2016.0459

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=5272391

Funtowicz, S. O., & Ravetz, J. R. (2018). Post-normal science. In *Companion to Environmental Studies*. Routledge.

Gardiner, S. M. (2010a). Ethics and climate change: An introduction. *WIREs Climate Change*, *1*(1), 54–66. https://doi.org/10.1002/wcc.16

Gardiner, S. M. (2010b). Is “Arming the Future” with Geoengineering Really the Lesser Evil? Some Doubts about the Ethics of Intentionally Manipulating the Climate System. In *Climate Ethics Essential Readings* (pp. 284–312).

Gardiner, S. M. (2011). *A Perfect Moral Storm: The Ethical Tragedy of Climate Change*. Oxford University Press.

Gardiner, S. M., & Fragnière, A. (2018). The Tollgate Principles for the Governance of Geoengineering: Moving Beyond the Oxford Principles to an Ethically More Robust Approach. *Ethics, Policy & Environment*, *21*(2), 143–174. https://doi.org/10.1080/21550085.2018.1509472

Gregory, R., Satterfield, T., & Hasell, A. (2016). Using decision pathway surveys to inform climate engineering policy choices. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(3), 560–565. https://www.jstor.org/stable/26467426

Grunwald, A. (2004). The normative basis of (health) technology assessment and the role of ethical expertise. *Poiesis & Praxis*, *2*(2), 175–193. https://doi.org/10.1007/s10202-003-0050-5

Grunwald, A. (2014). Technology Assessment for Responsible Innovation. In J. van den Hoven, N. Doorn, T. Swierstra, B.-J. Koops, & H. Romijn (Eds.), *Responsible Innovation 1: Innovative Solutions for Global Issues* (pp. 15–31). Springer Netherlands. https://doi.org/10.1007/978-94-017-8956-1_2

Grunwald, A. (2022). *Technikfolgenabschätzung: Einführung*. Nomos Verlag.

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=5272391

Grunwald, A., & Hillerbrand, R. (2021). *Handbuch Technikethik*. J. B. Metzler. https://link.springer.com/book/10.1007/978-3-476-04901-8

Guest, G., MacQueen, K. M., & Namey, E. E. (2012). *Applied Thematic Analysis*. SAGE Publications.

Guston, D. H. (2014). Understanding ‘anticipatory governance.’ *Social Studies of Science*, *44*(2), 218–242. https://doi.org/10.1177/0306312713508669

Guston, D. H., & Sarewitz, D. (2002). Real-time technology assessment. *Technology in Society*, *24*(1), 93–109. https://doi.org/10.1016/S0160-791X(01)00047-1

Hänsel, M. C., Drupp, M. A., Johansson, D. J. A., Nesje, F., Azar, C., Freeman, M. C., Groom, B., & Sterner, T. (2020). Climate economics support for the UN climate targets. *Nature Climate Change*, *10*(8), Article 8. https://doi.org/10.1038/s41558-020-0833-x

Harman, G. H. (1965). The Inference to the Best Explanation. *The Philosophical Review*, *74*(1), 88–95. https://doi.org/10.2307/2183532

Haszeldine, R. S., Flude, S., Johnson, G., & Scott, V. (2018). Negative emissions technologies and carbon capture and storage to achieve the Paris Agreement commitments. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *376*(2119), 20160447. https://doi.org/10.1098/rsta.2016.0447

Heyward, C. (2013). Situating and Abandoning Geoengineering: A Typology of Five Responses to Dangerous Climate Change. *PS: Political Science & Politics*, *46*(1), 23–27. https://doi.org/10.1017/S1049096512001436

Höfele, P. (2020). New technologies and the ‘heuristics of fear’. The meaning and prehistory of an emotion in Jonas, Heidegger and Hegel. *Hungarian Philosophical Review*, *64*, 166–182. http://filozofiaiszemle.net/2020/12/hungarian-philosophical-review-20201-

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=5272391

self-narrativity-emotions/

Höfele, P., Reuter, L., Estadieu, L., Livanec, S., Stumpf, M., & Kiesel, A. (2022). Connecting the methods of psychology and philosophy: Applying Cognitive-Affective Maps (CAMs) to identify ethical principles underlying the evaluation of bioinspired technologies. *Philosophical Psychology*, *0*(0), 1–24. https://doi.org/10.1080/09515089.2022.2113770

Hussain, Z., Binz, M., Mata, R., & Wulff, D. U. (2024). A tutorial on open-source large language models for behavioral science. *Behavior Research Methods*. https://doi.org/10.3758/s13428-024-02455-8

Hutto, C., & Gilbert, E. (2014). VADER: A Parsimonious Rule-Based Model for Sentiment Analysis of Social Media Text. *Proceedings of the International AAAI Conference on Web and Social Media*, *8*(1), Article 1. https://doi.org/10.1609/icwsm.v8i1.14550

Jacobs, J. (2002). *Dimensions of Moral Theory: An Introduction to Metaethics and Moral Psychology*. Blackwell Publishers.

Johansson, D. J. A., Azar, C., Lehtveer, M., & Peters, G. P. (2020). The role of negative carbon emissions in reaching the Paris climate targets: The impact of target formulation in integrated assessment models. *Environmental Research Letters*, *15*(12), 124024. https://doi.org/10.1088/1748-9326/abc3f0

Johnson, R. B., & Onwuegbuzie, A. J. (2004). Mixed Methods Research: A Research Paradigm Whose Time Has Come. *Educational Researcher*, *33*(7), 14–26. https://doi.org/10.3102/0013189X033007014

Jonas, H. (2020). *Das Prinzip Verantwortung: Versuch einer Ethik für die technologische Zivilisation*. Suhrkamp Verlag.

Keller, D. P., Feng, E. Y., & Oschlies, A. (2014). Potential climate engineering effectiveness and side effects during a high carbon dioxide-emission scenario. *Nature*

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=5272391

Communications, *5*(1), Article 1. https://doi.org/10.1038/ncomms4304

Keller, K., Helgeson, C., & Srikrishnan, V. (2021). Climate Risk Management. *Annual Review of Earth and Planetary Sciences*, *49*(1), 95–116. https://doi.org/10.1146/annurev-earth-080320-055847

Klepper, G., & Rickels, W. (2012). The Real Economics of Climate Engineering. *Economics Research International*, *2012*, e316564. https://doi.org/10.1155/2012/316564

Kluyver, T., Ragan-Kelley, B., Pérez, F., Granger, B., Bussonnier, M., Frederic, J., Kelley, K., Hamrick, J., Grout, J., Corlay, S., Ivanov, P., Avila, D., Abdalla, S., Willing, C., & Jupyter development team. (2016). *Jupyter Notebooks – a publishing format for reproducible computational workflows* (F. Loizides & B. Scmidt, Eds.; pp. 87–90). IOS Press. https://doi.org/10.3233/978-1-61499-649-1-87

Kuckartz, U., & Rädiker, S. (2022). *Qualitative Inhaltsanalyse. Methoden, Praxis, Computerunterstützung*. Beltz Juventa. https://content-select.com/de/portal/media/view/5e623532-20b8-4f33-b19e-4a1db0dd2d03?forceauth=1

Le Quéré, C., Peters, G. P., Friedlingstein, P., Andrew, R. M., Canadell, J. G., Davis, S. J., Jackson, R. B., & Jones, M. W. (2021). Fossil CO2 emissions in the post-COVID-19 era. *Nature Climate Change*, *11*(3), Article 3. https://doi.org/10.1038/s41558-021-01001-0

Lee, H., & Romero, J. (2023). *IPCC, 2023: Climate Change 2023: Synthesis Report. A Report of the Intergovernmental Panel on Climate Change. Contribution of Working Groups I, II and III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*. IPCC, Geneva, Switzerland. https://www.ipcc.ch/report/ar6/syr/downloads/report/IPCC_AR6_SYR_LongerReport.pdf

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=5272391

Lichtenstein, S., & Slovic, P. (2006). *The Construction of Preference*. Cambridge University Press.

Liu, P., Yuan, W., Fu, J., Jiang, Z., Hayashi, H., & Neubig, G. (2023). Pre-train, Prompt, and Predict: A Systematic Survey of Prompting Methods in Natural Language Processing. *ACM Computing Surveys*, *55*(9), 195:1-195:35. https://doi.org/10.1145/3560815

Low, S., Boettcher, M., Asayama, S., Baum, C., Borth, A., Brown, C., Clingerman, F., Dauvergne, P., De Pryck, K., Gupta, A., Honegger, M., Lenzi, D., Reitsma, R., Schenuit, F., Scott-Buechler, C., & Valenzuela, J. M. (2024). An earth system governance research agenda for carbon removal. *Earth System Governance*, *19*, 100204. https://doi.org/10.1016/j.esg.2024.100204

Lucivero, F. (2016). *Ethical Assessments of Emerging Technologies: Appraising the moral plausibility of technological visions*. Springer.

MacMartin, D. G., Irvine, P. J., Kravitz, B., & Horton, J. B. (2019). Technical characteristics of a solar geoengineering deployment and implications for governance. *Climate Policy*, *19*(10), 1325–1339. https://doi.org/10.1080/14693062.2019.1668347

Macnaghten, P., Davies, S. R., & Kearnes, M. (2019). Understanding Public Responses to Emerging Technologies: A Narrative Approach. *Journal of Environmental Policy & Planning*, *21*(5), 504–518. https://doi.org/10.1080/1523908X.2015.1053110

Marchau, V. A. W. J., Walker, W. E., Bloemen, P. J. T. M., & Popper, S. W. (Eds.). (2019). *Decision Making under Deep Uncertainty*. Springer Nature. https://doi.org/10.1007/978-3-030-05252-2

Masson-Delmotte, V., Zhai, P., Pörtner, H.-O., Roberts, D., Skea, J., Shukla, P. R., Pirani, A., Moufouma-Okia, W., Péan, C., Pidcock, R., Connors, S., Matthews, J. B. R., Chen, Y., Zhou, X., Gomis, M. I., Lonnoy, E., Maycock, T., Tignor, M., & Waterfield, T. (2018). *IPCC, 2018: Global Warming of 1.5°C. An IPCC Special Report on the*

Preprint not peer reviewed

impacts of global warming of 1.5°C above pre-industrial levels and related global greenhouse gas emission pathways, in the context of strengthening the global response to the threat of climate change, sustainable development, and efforts to eradicate poverty. Cambridge University Press. https://www.ipcc.ch/sr15/

Mayring, P. (2022). *Qualitative Content Analysis: A Step-by-Step Guide*. SAGE.

McCain, K., & Poston, T. (2017). *Best Explanations: New Essays on Inference to the Best Explanation*. Oxford University Press.

McLaren, D., Parkhill, K. A., Corner, A., Vaughan, N. E., & Pidgeon, N. F. (2016). Public conceptions of justice in climate engineering: Evidence from secondary analysis of public deliberation. *Global Environmental Change*, *41*, 64–73. https://doi.org/10.1016/j.gloenvcha.2016.09.002

McMillan, J. (2018). *The methods of bioethics: An essay in meta-bioethics* (First edition). Oxford University Press.

Meadows, D. H., Meadows, D. L., Randers, J., & Iii, W. W. B. (1972). *The Limits to Growth—Club of Rome*. Club of Rome. https://www.clubofrome.org/publication/the-limits-to-growth/

Merk, C., Pönitzsch, G., Kniebes, C., Rehdanz, K., & Schmidt, U. (2015). Exploring public perceptions of stratospheric sulfate injection. *Climatic Change*, *130*(2), 299–312. https://doi.org/10.1007/s10584-014-1317-7

Mirzadeh, I., Alizadeh, K., Shahrokhi, H., Tuzel, O., Bengio, S., & Farajtabar, M. (2024). *GSM-Symbolic: Understanding the Limitations of Mathematical Reasoning in Large Language Models* (No. arXiv:2410.05229). arXiv. https://doi.org/10.48550/arXiv.2410.05229

Möller, M., & Grießhammer, R. (2022). Prospective technology assessment in the Anthropocene: A transition toward a culture of sustainability. *The Anthropocene*

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=5272391

Review, *9*(2), 257–275. https://doi.org/10.1177/20530196221095700

Moosbrugger, H., & Kelava, A. (Eds.). (2020). *Testtheorie und Fragebogenkonstruktion*. Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-662-61532-4

Musschenga, B. (2009). Was ist empirische Ethik? *Ethik in der Medizin*, *21*(3), 187–199. https://doi.org/10.1007/s00481-009-0025-8

Muthén, L. K., & Muthen, B. O. (2017). *Mplus User's Guide. Eighth Edition.* [Computer software]. Muthén & Muthén. https://www.statmodel.com/html_ug.shtml

National Research Council. (2015). *Climate Intervention: Reflecting Sunlight to Cool Earth*. National Academies Press.

Neuber, F. (2018). *Buying Time with Climate Engineering? An analysis of the buying time framing in favor of climate engineering* [PhD Thesis, Karlsruher Institut für Technologie (KIT)]. https://doi.org/10.5445/IR/1000084294

Neuhäuser, C., Raters, M.-L., & Stoecker, R. (Eds.). (2023). *Handbuch Angewandte Ethik*. J.B. Metzler. https://doi.org/10.1007/978-3-476-05869-0

Newman, M. (2018). *Networks: An Introduction*. Oxford University Press.

Nordhaus, W. (1992). The 'Dice' Model: Background and Structure of a Dynamic Integrated Climate-Economy Model of the Economics of Global Warming. *Cowles Foundation Discussion Papers*. https://elischolar.library.yale.edu/cowles-discussion-paper-series/1252

Nordhaus, W. (2018). Projections and Uncertainties about Climate Change in an Era of Minimal Climate Policies. *American Economic Journal: Economic Policy*, *10*(3), 333–360. https://doi.org/10.1257/pol.20170046

Ott, K. (2011). Argumente für und wider „Climate Engineering ". In *Fallstudien zur Ethik in Wissenschaft, Wirtschaft, Technik und Gesellschaft* (pp. 198–210). KIT Scientific Publishing.

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=5272391

Ott, K. (2012). Domains of Climate Ethics. *Jahrbuch für Wissenschaft und Ethik*, *16*(1), 95–114. https://doi.org/10.1515/jfwe.2012.95

Ott, K., & Neuber, F. (2020). Climate engineering. In *Oxford Research Encyclopedia of Climate Science*. Oxford University Press.

Ouyang, L., Wu, J., Jiang, X., Almeida, D., Wainwright, C. L., Mishkin, P., Zhang, C., Agarwal, S., Slama, K., Ray, A., Schulman, J., Hilton, J., Kelton, F., Miller, L., Simens, M., Askell, A., Welinder, P., Christiano, P., Leike, J., & Lowe, R. (2022). *Training language models to follow instructions with human feedback* (No. arXiv:2203.02155). arXiv. https://doi.org/10.48550/arXiv.2203.02155

Pachauri, R. K., & Meyer, L. A. (2014). *IPCC, 2014: Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. https://www.ipcc.ch/report/ar5/syr/

Palm, E., & Hansson, S. O. (2006). The case for ethical technology assessment (eTA). *Technological Forecasting and Social Change*, *73*(5), 543–558. https://doi.org/10.1016/j.techfore.2005.06.002

Parkhill, K., & Pidgeon, N. (2011). Public Engagement on Geoengineering Research: Preliminary Report on the SPICE Deliberative Workshops. In *Public Engagement on Geoengineering Research*. https://eprints.whiterose.ac.uk/82892/

Parkhill, K., Pidgeon, N., Corner, A., & Vaughan, N. (2013). Deliberation and Responsible Innovation: A Geoengineering Case Study. In *Responsible Innovation* (pp. 219–239). John Wiley & Sons, Ltd. https://doi.org/10.1002/9781118551424.ch12

Paulo, N., & Bublitz, J. C. (2020). *Empirische Ethik: Grundlagentexte aus Psychologie und Philosophie*. Suhrkamp Verlag.

Pidgeon, N. (2021). Engaging publics about environmental and technology risks: Frames, values and deliberation. *Journal of Risk Research*, *24*(1), 28–46.

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=5272391

https://doi.org/10.1080/13669877.2020.1749118

Pidgeon, N., Parkhill, K., Corner, A., & Vaughan, N. (2013). Deliberating stratospheric aerosols for climate geoengineering and the SPICE project. *Nature Climate Change*, *3*(5), Article 5. https://doi.org/10.1038/nclimate1807

Pieper, A. (2017). *Einführung in die Ethik*. UTB.

Plazzotta, M., Séférian, R., Douville, H., Kravitz, B., & Tjiputra, J. (2018). Land Surface Cooling Induced by Sulfate Geoengineering Constrained by Major Volcanic Eruptions. *Geophysical Research Letters*, *45*(11), 5663–5671. https://doi.org/10.1029/2018GL077583

Pokorny, J. J., Norman, A., Zanesco, A. P., Bauer-Wu, S., Sahdra, B. K., & Saron, C. D. (2018). Network analysis for the visualization and analysis of qualitative data. *Psychological Methods*, *23*(1), 169–183. https://doi.org/10.1037/met0000129

Pörtner, H.-O., Roberts, D. C., Tignor, M., Poloczanska, E. S., Mintenbeck, K., Alegría, A., Craig, M., Langsdorf, S., Löschke, S., Möller, V., Okem, A., & Rama, B. (2022). *IPCC, 2022: Climate Change 2022: Impacts, Adaptation, and Vulnerability. Contribution of Working Group II to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press. https://www.ipcc.ch/report/sixth-assessment-report-working-group-ii/

Preston, C. J. (2012). *Engineering the Climate*. Rowman & Littlefield.

Preston, C. J. (2013). Ethics and geoengineering: Reviewing the moral issues raised by solar radiation management and carbon dioxide removal. *WIREs Climate Change*, *4*(1), 23–37. https://doi.org/10.1002/wcc.198

R Core Team. (2020). *R: A Language and Environment for Statistical Computing* [Computer software]. R Foundation for Statistical Computing. https://www.R-project.org/

Raschka, S. (2024). *Build a Large Language Model (From Scratch)*. Simon and Schuster.

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=5272391

Reuter, L., Mansell, J., Rhea, C., & Kiesel, A. (2022). Direct assessment of individual connotation and experience: An introduction to cognitive-affective mapping. *Politics and the Life Sciences*, *41*(1), 131–139. https://doi.org/10.1017/pls.2021.31

Reynolds, J. L., & Horton, J. B. (2020). An earth system governance perspective on solar geoengineering. *Earth System Governance*, *3*, 100043. https://doi.org/10.1016/j.esg.2020.100043

Rickels, W., Klepper, G., Dovern, J., Betz, G., Nadine, B., Güssow, K., Heintzenberg, J., Hiller, S., Hoose, C., Leisner, T., Oschlies, A., Platt, U., Proelß, A., Schäfer, S., Zürn, M., Cacean, S., & Renn, O. (2011). *Large-Scale Intentional Intervention s into the Climate System? Assessing the Climate Engineering Debate* [Scoping Report]. https://www.ifw-kiel.de/de/publikationen/books/large-scale-intentional-intervention-s-into-the-climate-system-assessing-the-climate-engineering-debate-6632/

Rockström, J., Steffen, W., Noone, K., Persson, Å., Chapin, F. S., Lambin, E. F., Lenton, T. M., Scheffer, M., Folke, C., Schellnhuber, H. J., Nykvist, B., de Wit, C. A., Hughes, T., van der Leeuw, S., Rodhe, H., Sörlin, S., Snyder, P. K., Costanza, R., Svedin, U., … Foley, J. A. (2009). A safe operating space for humanity. *Nature*, *461*(7263), Article 7263. https://doi.org/10.1038/461472a

Sand, M., Hofbauer, B. P., & Alleblas, J. (2023). Techno-fixing non-compliance—Geoengineering, ideal theory and residual responsibility. *Technology in Society*, *73*, 1–9. https://doi.org/10.1016/j.techsoc.2023.102236

Schwartz, M. S. (2016). Ethical Decision-Making Theory: An Integrated Approach. *Journal of Business Ethics*, *139*(4), 755–776. https://doi.org/10.1007/s10551-015-2886-8

Shepherd, J. G. (2009). *Geoengineering the climate: Science, governance and uncertainty* (p. 98) [Monograph]. Royal Society. https://doi.org/10/29)

Siegrist, M., & Árvai, J. (2020). Risk Perception: Reflections on 40 Years of Research. *Risk*

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=5272391

Analysis, *40*(S1), 2191–2206. https://doi.org/10.1111/risa.13599

Siegrist, M., & Hartmann, C. (2020). Consumer acceptance of novel food technologies. *Nature Food*, *1*(6), Article 6. https://doi.org/10.1038/s43016-020-0094-x

Sonntag, S., Ferrer González, M., Ilyina, T., Kracher, D., Nabel, J. E. M. S., Niemeier, U., Pongratz, J., Reick, C. H., & Schmidt, H. (2018). Quantifying and Comparing Effects of Climate Engineering Methods on the Earth System. *Earth's Future*, *6*(2), 149–168. https://doi.org/10.1002/2017EF000620

Sovacool, B. K., Baum, C. M., & Low, S. (2022). Determining our climate policy future: Expert opinions about negative emissions and solar radiation management pathways. *Mitigation and Adaptation Strategies for Global Change*, *27*(8), 58. https://doi.org/10.1007/s11027-022-10030-9

Sovacool, B. K., Baum, C. M., & Low, S. (2023). Beyond climate stabilization: Exploring the perceived sociotechnical co-impacts of carbon removal and solar geoengineering. *Ecological Economics*, *204*, 107648. https://doi.org/10.1016/j.ecolecon.2022.107648

Steegen, S., Tuerlinckx, F., Gelman, A., & Vanpaemel, W. (2016). Increasing Transparency Through a Multiverse Analysis. *Perspectives on Psychological Science*, *11*(5), 702–712. https://doi.org/10.1177/1745691616658637

Sühr, T., Dorner, F. E., Samadi, S., & Kelava, A. (2024). *Challenging the Validity of Personality Tests for Large Language Models* (No. arXiv:2311.05297). arXiv. https://doi.org/10.48550/arXiv.2311.05297

ter Mors, E., Terwel, B. W., Daamen, D. D. L., Reiner, D. M., Schumann, D., Anghel, S., Boulouta, I., Cismaru, D. M., Constantin, C., de Jager, C. C. H., Dudu, A., Esken, A., Falup, O. C., Firth, R. M., Gemeni, V., Hendriks, C., Ivan, L., Koukouzas, N., Markos, A., … Ziogou, F. (2013). A comparison of techniques used to collect informed public opinions about CCS: Opinion quality after focus group discussions

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=5272391

versus information-choice questionnaires. *International Journal of Greenhouse Gas Control*, *18*, 256–263. https://doi.org/10.1016/j.ijggc.2013.07.015

Thagard, P. (1998). Ethical coherence. *Philosophical Psychology*, *11*(4), 405–422. https://doi.org/10.1080/09515089808573270

Thagard, P. (2000). *Coherence in Thought and Action*. MIT Press.

Thagard, P. (2010). EMPATHICA: A Computer Support System with Visual Representations for Cognitive-Affective Mapping. *Workshops at the Twenty-Fourth AAAI Conference on Artificial Intelligence*, 79–81. https://www.aaai.org/ocs/index.php/WS/AAAIW10/paper/view/1981

Thomas, G., Pidgeon, N., & Roberts, E. (2018). Ambivalence, naturalness and normality in public perceptions of carbon capture and storage in biomass, fossil energy, and industrial applications in the United Kingdom. *Energy Research & Social Science*, *46*, 1–9. https://doi.org/10.1016/j.erss.2018.06.007

Touvron, H., Lavril, T., Izacard, G., Martinet, X., Lachaux, M.-A., Lacroix, T., Rozière, B., Goyal, N., Hambro, E., Azhar, F., Rodriguez, A., Joulin, A., Grave, E., & Lample, G. (2023). *LLaMA: Open and Efficient Foundation Language Models* (No. arXiv:2302.13971). arXiv. https://doi.org/10.48550/arXiv.2302.13971

Tunstall, L., Werra, L. von, & Wolf, T. (2022). *Natural Language Processing with Transformers*. O'Reilly Media, Inc.

United Nations Environment Programme. (1992). *Report of the United Nations Conference on Environment and Development*. https://www.un.org/en/conferences/environment/rio1992

Van Rossum, G., & Drake, F. L. (2009). *Python 3 Reference Manual*. CreateSpace.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł. ukasz, & Polosukhin, I. (2017). Attention is All you Need. *Advances in Neural*

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=5272391

Information Processing Systems, *30*. https://proceedings.neurips.cc/paper/2017/hash/3f5ee243547dee91fbd053c1c4a845aa-Abstract.html

Vink, S., van Tartwijk, J., Verloop, N., Gosselink, M., Driessen, E., & Bolk, J. (2016). The articulation of integration of clinical and basic sciences in concept maps: Differences between experienced and resident groups. *Advances in Health Sciences Education*, *21*(3), 643–657. https://doi.org/10.1007/s10459-015-9657-2

Wang, Y., Ma, X., Zhang, G., Ni, Y., Chandra, A., Guo, S., Ren, W., Arulraj, A., He, X., Jiang, Z., Li, T., Ku, M., Wang, K., Zhuang, A., Fan, R., Yue, X., & Chen, W. (2024). *MMLU-Pro: A More Robust and Challenging Multi-Task Language Understanding Benchmark (Published at NeurIPS 2024 Track Datasets and Benchmarks)* (No. arXiv:2406.01574). arXiv. https://doi.org/10.48550/arXiv.2406.01574

Welsby, D., Price, J., Pye, S., & Ekins, P. (2021). Unextractable fossil fuels in a 1.5 °C world. *Nature*, *597*(7875), Article 7875. https://doi.org/10.1038/s41586-021-03821-8

White, J., Fu, Q., Hays, S., Sandborn, M., Olea, C., Gilbert, H., Elnashar, A., Spencer-Smith, J., & Schmidt, D. C. (2023). *A Prompt Pattern Catalog to Enhance Prompt Engineering with ChatGPT* (No. arXiv:2302.11382). arXiv. https://doi.org/10.48550/arXiv.2302.11382

Wibeck, V., Hansson, A., Anshelm, J., Asayama, S., Dilling, L., Feetham, P. M., Hauser, R., Ishii, A., & Sugiyama, M. (2017). Making sense of climate engineering: A focus group study of lay publics in four countries. *Climatic Change*, *145*(1), 1–14. https://doi.org/10.1007/s10584-017-2067-0

Workman, M., Dooley, K., Lomax, G., Maltby, J., & Darch, G. (2020). Decision making in contexts of deep uncertainty—An alternative approach for long-term climate policy. *Environmental Science & Policy*, *103*, 77–84.

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=5272391

https://doi.org/10.1016/j.envsci.2019.10.002

Xie, Y., Allaire, J. J., & Grolemund, G. (2018). *R Markdown: The Definitive Guide*. Chapman and Hall/CRC. https://doi.org/10.1201/9781138359444

Yan, L., Sha, L., Zhao, L., Li, Y., Martinez-Maldonado, R., Chen, G., Li, X., Jin, Y., & Gašević, D. (2024). Practical and ethical challenges of large language models in education: A systematic scoping review. *British Journal of Educational Technology*, *55*(1), 90–112. https://doi.org/10.1111/bjet.13370

Yang, J., Jin, H., Tang, R., Han, X., Feng, Q., Jiang, H., Zhong, S., Yin, B., & Hu, X. (2024). Harnessing the Power of LLMs in Practice: A Survey on ChatGPT and Beyond. *ACM Trans. Knowl. Discov. Data*, *18*(6), 160:1-160:32. https://doi.org/10.1145/3649506

Zhang, H., Wang, F., Li, J., Duan, Y., Zhu, C., & He, J. (2022). Potential Impact of Tonga Volcano Eruption on Global Mean Surface Air Temperature. *Journal of Meteorological Research*, *36*(1), 1–5. https://doi.org/10.1007/s13351-022-2013-6

Declaration of generative AI and AI-assisted technologies in the writing process

During the preparation of this work, the authors utilized one Llama model, which were developed by Meta to analyze data and generate textual summaries. Specifically, the AI model "Llama-3.1-70B-Instruct" was employed for generating structured summaries and comparative syntheses of qualitative data based on predefined prompts. These prompts guided the model to extract and organize key ethical arguments from coded datasets, and to compare layperson interpretations with expert definitions. After using this tool, the authors thoroughly reviewed and edited the content to ensure accuracy, coherence, and alignment with the study's objectives. The authors take full responsibility for the final content of the published article.

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=5272391

Appendix A: Sample statistics Cognitive-Affective Maps

The subsequent report was generated by utilizing the "Get Report" function within the Data Analysis Tool (Fenn et al., under review).

Description of dataset

In total, we collected 58 CAMs, of which 0 (0%) CAMs were excluded from further analysis. Participants drew on average 25.4 (*SD* = 2.06) concepts (whereby 34% were positive, 46% negative, 12% neutral and 8% ambivalent). Please note that the technical settings required participants to draw at least 24 concepts. On average, 44.21 (*SD* = 32.49) connectors were drawn. 82% of the connectors were agreeing and 18% disagreeing. Furthermore, 21% of the connectors were bidirectional and 79% unidirectional. The valence for the concepts range from [-3,-1] for negative and [1,3] for positive concepts, with ambivalent and neutral concepts being assigned a value of 0. The mean average valence over all the CAMs was -0.33 (*SD* = 0.51). In 14% of the non-deleted CAMs one or more of the predefined concepts were removed by the participants.

Summarizing concepts

We summarized the CAMs using the dedicated Data Analysis Tool. The Data Analysis Tool generates a protocol which tracks each summarizing step so that the summarizing process is completely transparent. The 1,063 raw unique concepts (1,473 in total) were summarized to 41 concepts using 101 times the "approximate matching", 203 times the "searching terms", 4 times the "search for synonyms" and 10 times the "apply word2vec model" functionalities.

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=5272391

Appendix B: Summary of Ethical Arguments in Cognitive-Affective Maps

In table SM2, the frequencies and average valence of all summarized concepts to ethical arguments are presented (cf., Tab. 1), whereby an overall searchable wordlist can be found on OSF (https://osf.io/rhxfn). In the table, the last row "mimics nature" is not an ethical argument but could indicate the above discussed "Natural-is-better" heuristic (see section 2.1 in main article). For the meaning of single variables please see "Note" below the table.

Table SM2.

Frequencies of all summarized concepts to ethical arguments

Concept	*N*	*M*	*SD*	Examples
Side-effects not predictable	264	-2.25	0.98	uncertain whether it would contribute to more acid rain; unknown side effects
Greater Good	143	2.22	0.9	environmental benefits; improve global health
Risk of Governance	63	-1.86	1.05	government focuses policy on electoral cycles [..]
Moral Hazard	37	-1.46	1.76	allows society to carry on polluting the planet
Risk Transfer to the Future	37	-2.32	0.82	social or political conflicts over the use; cause war?
Informed Consent	31	-1.32	1.19	impossible to get everyone to agree on it
Long-Term Control	24	-1.92	0.97	could be hard to get all countries to commit for a long period of time
Technological Fix	21	0.48	1.97	a good solution [..]; less human burden
Termination Problem	19	-1.84	1.17	if stopped temperatures would rise rapidly
Unfair distribution of effects and power	17	-1.41	1.12	might affect one region more than another, causing tensions
Arming the Future	14	2	1.11	creating a better future for future generations
Maintaining the Status Quo	12	-1.83	1.11	companies, governments will carry on polluting and using bad technology/fossil fuels
Buying Time	11	1	0.77	allows for more time to create long term solutions
Lesser-evil	8	-0.62	1.19	emergency solution
Risk of Unilateral Use	8	-2.75	0.46	governments in wealthier nations would be able to exert control over poorer nations
Betrayal of Divine Creation	6	-2.33	0.82	Is SAI another way in which humans mess things up by playing God?
Do it Alone	4	1.5	1.29	can be implemented by a wide variety of nations
Dual Use	3	-3	0	governments could weaponise [..] the use
Hubris Argument	2	-1	1.41	give us the illusion we can carry on as we have been
mimics nature	10	1.3	1.16	Technology that comes from [..] volcanos

Note. N is the total frequency of summarized concepts to ethical argument, *M* the mean valence, *SD* the standard deviation of the mean valence, and in the Examples column are typical examples from the text or comments of the drawn concepts.

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=5272391

Appendix C: Summary of Assigned Ethical Arguments in Open Text

In table SM3, the frequencies and percentages of all assigned ethical arguments are presented. In total, only 553 ethical arguments were considered (see main article). Raters could assign multiple ethical arguments to the text answer of a single participant (for meaning of single variables, see “Note”).

Table SM3.
Frequencies of all assigned ethical arguments

Ethical argument	*N*	percentage	number of co-occurrence
Emergency Case	202	23.01	160
Informed Consent	126	14.35	153
Lesser-evil	119	13.55	94
99 (residual category)	87	9.91	87
Side-effects unseen / not predictable	74	8.43	83
Greater Good	71	8.09	38
Technological Fix	31	3.53	36
Unfair distribution of effects and power	26	2.96	50
Moral Hazard	24	2.73	33
Termination Problem	18	2.05	30
Arming the Future	17	1.94	14
Hubris Argument	15	1.71	16
Buying Time	14	1.59	18
Risk of Unilateral Use	12	1.37	21
Maintaining the Status Quo	10	1.14	16
Risk of Governance	10	1.14	17
Betrayal of Divine Creation	7	0.80	11
Long-Term Control	6	0.68	12
Risk Transfer to the Future	4	0.46	8
Do it Alone	3	0.34	4
Unstoppable Deployment if researched	2	0.23	7
Dual Use	0	0	0

Note. N is the frequency a single ethical argument was assigned by all raters, percentage is the respective percentage (divided by total sum of number) and the “number of co-occurrence” indicates how often a single ethical argument was assigned together with all the other ethical arguments.

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=5272391

Appendix D: Introduction to Large Language Models

LLMs are advanced artificial intelligence tools that excel at processing and generating human-like text, which is shown by benchmark testing (e.g., Chiang et al., 2024; Wang et al., 2024). These models are trained on extensive datasets comprising billions to trillions of tokens - fundamental units of text processed by LLMs, which may represent entire words, subwords, or characters, depending on the tokenizer - enabling them to identify patterns, relationships, and structures inherent in language (Caelen & Blete, 2023; Tunstall et al., 2022). A defining characteristic of LLMs is their ability to predict the next word or token in a sequence. For example, when prompted with “The capital of France is”, an LLM will likely predict “Paris,” leveraging probabilistic patterns learned from its training data (Hussain et al., 2024; Vaswani et al., 2017)

LLMs are built on the generative pretrained transformer (GPT) architecture, which relies on self-attention mechanisms to effectively process input text. Self-attention enables the model to focus on the most relevant parts of the input sequence, capturing contextual meaning at varying scales (Vaswani et al., 2017). This architecture excels across a wide array of tasks, including summarization, machine translation, classification, and creative writing (Dubey et al., 2024; Hussain et al., 2024; Touvron et al., 2023). The training of an LLM typically involves two stages: pre-training and fine-tuning. During pre-training, the model is exposed to a vast corpora, including books, scientific articles, and internet-sourced text, to learn general language patterns resulting in a foundational model (Raschka, 2024). Fine-tuning subsequently adapts the pre-trained model to specific tasks or domains using curated datasets and human feedback (Brown et al., 2020; Ouyang et al., 2022). Fine-tuned models often include the term “Instruct” in their name to denote alignment with task-specific objectives, as seen in the Llama-3.1-70B-Instruct model (Dubey et al., 2024). LLMs are versatile tools with applications spanning multiple domains. They can condense large datasets

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=5272391

into concise, interpretable summaries, classify text into predefined categories, generate synthetic data for machine learning applications, and assist researchers in identifying themes and patterns within complex datasets (Debelak et al., 2024; Hussain et al., 2024; Yang et al., 2024).

Despite their transformative potential, LLMs have inherent limitations. They excel at pattern recognition and text generation but lack genuine understanding, reasoning, or cognitive abilities akin to humans (Mirzadeh et al., 2024). Their reliance on training data makes them susceptible to perpetuating biases or inaccuracies present in the underlying datasets (Sühr et al., 2024; Yan et al., 2024). Furthermore, their context length - the number of tokens that can be processed simultaneously - varies between models, ranging from 8,000 tokens to over 100,000 tokens, which can restrict their utility for lengthy or complex tasks (Dubey et al., 2024; Tunstall et al., 2022).

Effective use of LLMs requires adherence to well-established prompting strategies. Structured prompts, which include clear instructions, relevant context, and examples of desired outputs, can significantly enhance model performance (Liu et al., 2023; White et al., 2023).

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=5272391

A Data-driven Approach for Identifying Ethical Concerns of Climate Engineering Technologies - Stratospheric Aerosol Injection as a use-case

Julius Fenn[1], Michael Gorki[1, 2], Philipp Höfele[3], Louisa Estadieu[2,4], Christophe Becht[1], Lars Kulbe[1, 2], Johannes Gekeler[1] and Andrea Kiesel[1, 2]

[1]Institute of Psychology, University of Freiburg, Germany

[2]Cluster of Excellence livMatS @ FIT Freiburg Center for Interactive Materials and Bioinspired Technologies, University of Freiburg, Germany

[3]Young Academy for Sustainability Research (YAS) at Freiburg Institute for Advanced Studies (FRIAS), University of Freiburg, Germany

[4] Department of Humanities, Social and Political Sciences, ETH Zürich, Switzerland

Author Notes

Julius Fenn https://orcid.org/0000-0002-7569-8648

E-Mail: julius.fenn@psychologie.uni-freiburg.de

Michael Gorki https://orcid.org/0009-0007-6138-4043

E-Mail: michael.gorki@livmats.uni-freiburg.de

Philipp Höfele https://orcid.org/0000-0002-8682-9965

E-Mail: philipp.hoefele@frias.uni-freiburg.de

Louisa Estadieu https://orcid.org/0000-0002-3164-4445

E-Mail: louisa.estadieu@ethz.ch

Christophe Becht https://orcid.org/0009-0005-8483-2588

E-Mail: christophe.becht@psychologie.uni-freiburg.de

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=5272391

Lars Kulbe https://orcid.org/0009-0005-6244-6002

E-Mail: lars.kulbe@student.uva.nl

Johannes Gekeler https://orcid.org/0009-0007-4936-5632

E-Mail: johannes.gekeler@students.uni-freiburg.de

Andrea Kiesel https://orcid.org/0000-0001-5564-010X

E-Mail: kiesel@psychologie.uni-freiburg.de

Lars Kulbe is now at the Department of Psychology, University of Amsterdam.

*corresponding author: Julius Fenn, Department of Psychology, University of Freiburg, Engelbergerstraße 41, 79085 Freiburg. julius.fenn@psychologie.uni-freiburg.de

Funding

The work leading to this publication was supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy – EXC-2193/1 – 390951807, as well as by the Eva Mayr-Stihl Foundation as part of the "Young Academy for Sustainability Research" at the "Freiburg Institute for Advanced Studies" (FRIAS), University of Freiburg.

Acknowledgement

This work was inspired by discussions with our colleagues from the Living, Adaptive and Energy-autonomous Materials Systems excellence cluster of the University of Freiburg. Many thanks to my co-author Philipp Höfele for making me aware of the importance of an empirically informed ethics.

Declaration of Interest statement

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=5272391

Abstract

Limiting global warming to 1.5 °C has intensified interest in climate engineering technologies such as Stratospheric Aerosol Injection (SAI), which mimic volcanic cooling. Given potential insufficiency of mitigation alone, ethical examination of SAI is imperative. This study investigates whether laypersons' ethical reasoning about SAI can be empirically identified. Using a multi-method design, we combined Cognitive-Affective Maps (CAMs) and open-ended textual responses to elicit twenty ethical concerns. Large Language Models (LLMs) synthesized lay perspectives and compared them against formal ethical definitions. Results revealed diverse ethical considerations, including governance, risk, equitable deployment, and emergency use. In contrast to formal definitions, lay participants foregrounded practical implications, social trust, and personal experience. Our findings demonstrate the utility of integrating data sources for empirical ethics research and underscore the complexity of public ethical discourse on SAI. This approach promotes more inclusive, evidence-based dialogue on the responsible development and governance of climate engineering technologies.

Keywords: climate engineering; climate change; Cognitive-Affective Maps; qualitative content analysis; network analysis; large language models

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=5272391

1. Introduction

Current attempts to lower greenhouse gas emissions and to adapt to the effects of climate change are still insufficient according to multiple authors (e.g., Le Quéré et al., 2021; Lee & Romero, 2023; Pörtner et al., 2022; Welsby et al., 2021) or as indicated by recent reports of the Climate Action Tracker[1]. The most recent report of the Intergovernmental Panel on Climate Change (IPCC, AR6 Synthesis Report from 2023) warns that current mitigation contributions "make it likely that warming will exceed 1.5°C during the 21st century" (Lee & Romero, 2023, p. 23). This concern is underscored by the fact that global average temperatures surpassed the 1.5°C threshold above pre-industrial levels for the first time in 2024[2]. In the perspective of the environmental philosopher Gardiner (Gardiner, 2011) this marks an "environmental tragedy" – despite (scientific) facts on climate change being well known (e.g., in The Limits to Growth report by Meadows et al., 1972), current actions are not sufficiently effective.

To increase the chances that temperature increase is limited to 1.5°C, multiple reports and scientific studies emphasize that Climate Engineering Technologies (CETs), especially "negative emissions" technologies, are necessary (e.g., Anderson & Peters, 2016; Haszeldine et al., 2018; Johansson et al., 2020; Welsby et al., 2021). In general, there are two distinct approaches of CETs to address climate change (see Caviezel & Revermann, 2014; Dowling, 2018; Heyward, 2013; National Research Council, 2015; Shepherd, 2009) with varying ethical concerns (see Betz & Cacean, 2012; Ott & Neuber, 2020; Rickels et al., 2011). Carbon Dioxide Removal technologies, also called "negative emissions" technologies,

[1] Climate Action Tracker is an independent scientific project that analyzes data from 39 countries, collectively covering 85% of global emissions, to produce its reports; see Climate Action Tracker Thermometer: https://climateactiontracker.org/global/cat-thermometer/

[2] See "Copernicus: 2024 is the first year to exceed 1.5°C above pre-industrial level": https://climate.copernicus.eu/copernicus-2024-first-year-exceed-15degc-above-pre-industrial-level; The Copernicus Climate Change Service (C3S), operated by the European Centre for Medium-Range Weather Forecasts (ECMWF) as part of the EU's Copernicus programme, provides free, reliable, and up-to-date data on climate and environmental changes.

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=5272391

remove carbon dioxide from the atmosphere, which addresses the root cause of climate change. In contrast, Solar Radiation Management technologies seek to reflect a small percentage of solar radiation back into space before it reaches the earth. Such technologies are already included in most scenarios (so-called "Integrated Assessment Models") of the IPCC reports, which quantitatively describe key human and earth system processes of climate change (e.g., Lee & Romero, 2023; Masson-Delmotte et al., 2018; Pachauri & Meyer, 2014; Pörtner et al., 2022).

According to Sand et al. (2023), CETs can be framed as a "techno-fix" for the problem of insufficient climate mitigation. Such CETs do not demand behavioral changes of people and might be implemented more easily and faster than large societal transformations (Preston, 2012, 2013). Because such technologies could free ourselves from the obligation to reduce emissions and thereby impact our moral agency (Gardiner, 2010a), framing CETs as a "techno-fix" is therefore highly contested from an ethical standpoint (Corner & Pidgeon, 2014). Due to, for example, possible unknown side-effects, the problem of climate change could even be enlarged (see expert interviews in Sovacool et al., 2022, 2023).

Given the critical role of CETs in addressing climate issues, it is critical to empirically investigate their ethical concerns associated with their development and implementation. Such inquiry could finally support the responsible and informed governance of these emerging technologies (Low et al., 2024; Reynolds & Horton, 2020). To this end, we focus on a specific Solar Radiation Management technology known as Stratospheric Aerosol Injection (SAI) as a use case, yet our proposed methodology can be easily adjusted and applied to different types of emerging CETs. Investigating ethical concerns of SAI is crucial because the technology is highly efficient in comparison to other CETs (D. P. Keller et al., 2014; Sonntag et al., 2018), timely and relatively cheap (Shepherd, 2009). SAI can decrease the amount of incoming solar radiation by releasing sulfur particles into the stratosphere,

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=5272391

enhancing the aerosol layer’s reflective properties. This technology, most prominently proposed by Crutzen (2006), mimics the natural cooling effect observed after volcanic eruptions (e.g., Mount Tambora in 1815 or Mount Pinatubo in 1991), during which sulfur particles are released into the atmosphere (cf., Plazzotta et al., 2018; Zhang et al., 2022). SAI could be deployed in an emergency case when mitigation efforts have been insufficient. However, there are fundamental ethical concerns (for an overview, see Tab. 1 in section 1.2.), which lead some scientists to advocate a Non-Use Agreement (Biermann et al., 2022). For example, even the act of just researching SAI could by itself decrease the motivation of individuals and governments to implement necessary, far-reaching mitigation policies (ethical argument of “Moral Hazard”) and transfer the risks of climate change to future generations, thus putting them in a dilemma to finally deploy SRM technologies, which is the ethical argument of “Risk Transfer to the Future” (e.g., Callies, 2019; Preston, 2012).

We propose a methodology that relates two heterogeneous data sources - Cognitive-Affective Maps (CAMs) and open text - and apply three types of data analyses: network analyses, qualitative content analysis, and Large Language Models (LLMs) to answer our main research question: Is it possible to empirically identify ethical arguments of laypersons regarding our use-case SAI? The article is organized as follows: In Section 1.1, we briefly motivate the need for empirically informed ethics, followed by a discussion of individual ethical arguments in Section 1.2. Section 2 describes the overall study design, which includes two time points for data collection and two different types of data. Section 3 presents the results for both data sources, along with their respective statistical procedures. Finally, Section 4 provides an overview of all results and proposes future research questions.

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=5272391

1.1. Motivation of empirical informed ethics

The 15th Principle of the Rio Declaration on Environment and Development, the so-called Precautionary Principle, states that if "there are threats of serious or irreversible damage, lack of full scientific certainty shall not be used as a reason for postponing cost-effective measures to prevent environmental degradation" (United Nations Environment Programme, 1992, p. 6). Thus, it could be argued that CETs, which are at an early development stage, should be evaluated regarding benefits and costs while also considering risks which are not of "full scientific certainty" (cf., Buckley et al., 2017). To find an optimal balance between the economic costs of greenhouse gas reductions (e.g., reduced consumption) and their benefits (e.g., reduced gross damages) based on the Dynamic Integrated Climate-Economy Model by Nordhaus (1992), multiple climate scenarios have been simulated (e.g., De Bruin et al., 2009; Johansson et al., 2020; Nordhaus, 2018). There is a vivid discussion in the economic literature, for instance, on whether CETs are relatively inexpensive compared to mitigation (see Barrett, 2008 vs. Klepper & Rickels, 2012) or on how the parameters of such models should be adjusted to account for factors such as intergenerational welfare (Hänsel et al., 2020).

Relying exclusively on expert-driven strategies or climate scenario modeling for decision-making in addressing climate change is problematic due to deep uncertainties inherent in the dynamics of the climate system and the complexity of the planetary boundaries we are approaching (e.g., Rockström et al., 2009). Neither is there expert consensus on what outcomes (e.g., cost-effectiveness vs. intergenerational equity) should be aimed for, nor which climate policies should be pursued (K. Keller et al., 2021; Marchau et al., 2019; Workman et al., 2020). For CETs specifically, irreducible uncertainties persist, such as those arising from the inherent complexities of the Earth system, human error, and limitations in predictive models (e.g., Betz & Cacean, 2012; Neuber, 2018; Rickels et al.,

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=5272391

2011), underscoring the need for a nuanced and multi-perspective approach to their evaluation. We argue that it is at least central to ethically evaluate emerging CETs in real time to decide between a cautious (more conservative) approach, linked to the principle of precaution (cf., Höfele, 2020; Jonas, 2020), and a constructive (liberal) approach, linked to an optimistic principle of innovation (cf. Bindé, 2000; Grunwald, 2014; Guston & Sarewitz, 2002; Musschenga, 2009).

To guide decision-making processes under conditions of such high system uncertainty and high decision stakes, Funtowicz and Ravetz (2018) emphasize the necessity of including “extended peer communities”. Standard expert-analytical assessments alone are inadequate in such contexts, which are emblematic of “post-normal science” (cf., Workman et al., 2020), where the complexity and stakes of decisions demand broader participation and diverse perspectives. Incorporating laypersons’ perspectives alongside expert-driven approaches not only makes decisions more attuned to diverse values and ethical concerns, generating new substantive insights, but also enhances the legitimacy of decisions and fosters greater trust in policymaking processes (Fiorino, 1990; Pidgeon, 2021; Wibeck et al., 2017).

The inclusion of laypersons could even foster structures of *anticipatory governance*, which is the capacity “extended through society that can act on a variety of inputs to manage emerging knowledge-based technologies while such management is still possible” (Guston, 2014, p. 219). It is also possible, for example by using robust decision making, to identify climate policy options which are robust in many possible future scenarios, whereby the process of making decisions adapts to the unfolding future (Marchau et al., 2019). In the context of CETs, considering the (ethical) concerns of all stakeholders affected by such technologies would, at best, lead to changes in the research and implementation process (cf., Frumhoff & Stephens, 2018; Gardiner & Fragnière, 2018). In our opinion, such an empirically informed ethics, i.e. the integration of laypersons' perspectives at the early stages

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=5272391

of CET development, could enable the identification of potential ethical challenges and foster adaptive, socially responsive approaches to governance.

In the next section, we present an overview of ethical arguments regarding CETs already identified in the literature. These ethical arguments are applied to structure the view of laypersons and relate their answers (see CAMs, open text sections) to ethically established arguments (for a similar procedure see Höfele et al., 2022).

1.2. Ethical Arguments

CETs and emerging technologies in general are accompanied by uncertainties that call for normative regulations. Thereby, the appropriate use and design of technologies, as well as the acceptable consequences associated with them, are unknown (cf., Grunwald, 2004, 2022). Such uncertainties are especially dealt with by means of two interrelated ethical perspectives (see Cotton, 2014; Grunwald & Hillerbrand, 2021; Pieper, 2017). *Normative ethics* investigates the criteria for determining the moral rightness or wrongness of actions and virtues. *Applied ethics* involves applying ethical principles or theories to specific problems and conflicts in various life areas. It has developed into several independent subfields, including, for example, medical, environmental, animal, science and technology, political, legal, professional, and business ethics, which have expanded significantly over the past 20 years (Grunwald & Hillerbrand, 2021; Neuhäuser et al., 2023). *Empirical ethics* goes a step further by integrating empirical data to examine how moral values are understood and lived in practice, enriching the normative analysis with insights from real-world behaviors and attitudes (Paulo & Bublitz, 2020). Ethical arguments related to SAI belonging to the realm of applied ethics often take a deductive form (argument is deductively valid if its conclusion logically follows from the premises) and make use of descriptive empirical and normative premises (cf., Betz & Cacean, 2012; Neuber, 2018). For example, the "Lesser-Evil" argument

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=5272391

states that deploying SAI is necessary to prevent catastrophic global warming, and includes a descriptive premise ("[a]t some future point in time t, we may end up in a situation where [...] the worst possible impacts of the deployment of the CET are clearly less severe than the worst possible consequences of not deploying it"; Betz & Cacean, 2012, p. 32) and a normative premise ("one should choose the option for action with the comparatively best worst possible consequences"; Betz & Cacean, 2012, p. 32). Due to the reliance on often changing descriptive premises and their complexities, ethical arguments often cannot be definitively justified as true or right, in our opinion. Therefore, they need to be continually critically evaluated to identify the best evidence (descriptive premises) for a certain conclusion (e.g., deploying SAI is the best option in a specific future context). Such a procedure is closely linked to the theory of "The Inference to the Best Explanation" (Harman, 1965; McCain & Poston, 2017), as such an ethical argument "includes relevant considerations that give us reason for thinking that the conclusion is likely to follow" (McMillan, 2018, p. 113).

The ethical arguments regarding CETs in Table 1, which are applied in the following sections, are based on multiple reports and scientific articles from authors in the field of philosophy and ethics (cf., Betz & Cacean, 2012; Neuber, 2018; Ott, 2011, 2012; Ott & Neuber, 2020; Preston, 2012, 2013; Rickels et al., 2011) as well as from authors in the field of the social sciences, whereby we only considered qualitative studies investigating the ethical concerns of laypersons (Carr & Yung, 2018; Corner et al., 2011, 2013; McLaren et al., 2016; Parkhill et al., 2013; Parkhill & Pidgeon, 2011; Pidgeon et al., 2013; Wibeck et al., 2017). We iteratively generated and adjusted definitions of the ethical arguments during team discussions.

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=5272391

Table 1

Overview of identified ethical arguments regarding CETs

Ethical Argument	**Definition**	**Coding Rules**	**General Evaluation**
Moral Hazard (also: "Undermining Better Options")	• Researching and developing CETs may foster the idea of a technical climate solution, which might reduce people's enthusiasm for pursuing (potentially challenging) mitigation measures / mitigation policies • Solely investing in CETs research and development may divert resources from mitigation efforts • Lobby groups and media hype around CETs could further undermine emission abatement and adaptation measures	Compared to "Arming the Future" negative future perspective on the research of CETs. Compared to "Risk Transfer to the Future" the argument focuses on mitigation efforts / policies (no global perspective). Respondents do not need to emphasize the last bullet point.	Negative
Risk Transfer to the Future	• Research and development of CETs transfers risks to future generations • CETs can create new conflicts and may trigger wars • Deciding to deploy or not deploy these technologies will likely lead to future dilemmas	Compared to "Arming the Future" negative future perspective on the research of CETs. Compared to "Moral Hazard" this argument takes a more global perspective (e.g., "future generation").	Negative
Arming the Future	• There is a moral obligation to consider all options for future generations • Available CETs give future generations the ability to control the climate • Future generations should have the freedom to choose whether to use CETs	Compared to "Moral Hazard" and "Risk Transfer to the Future" positive future perspective on the research of CETs.	Positive
Technological Fix	• Technological fixes are attractive when citizens fail to make necessary behavioral changes • They are often simpler, faster, and require less effort than extensive social transformations • However, such solutions tinker with symptoms instead of resolving the causes, because it would permit continuing high levels of consumption, waste, and greenhouse gas emissions	Remark: ambivalent argument, depending if a respondent perceives a "Technological Fix" as something positive or negative (last bullet point).	Ambivalent

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=5272391

Table 1 (continued).

Maintaining the Status Quo	• CETs are a "pseudo-solution" that maintains the status quo and benefits industrial sectors and business branches that are the most reactionary in terms of climate policy • If CETs are controlled by big business, it may even reinforce the status quo • There is suspicion around the motivations, benefits, and secrecy of industries developing CETs	Respondents need to highlight in any form the “status quo”, which is perceived negatively.	Negative
Unstoppable Deployment if researched	• CETs research may generate internal momentum for deployment, even if unnecessary or not desirable • capital-intensive CETs would only be recouped over a long period of time • more investment in CETs research makes it harder to prevent future deployment	Compared to “Maintaining the Status Quo” this ethical argument emphasizes a *path dependency*, which are past decisions, which influence the choices and development of a system, often leading it down a specific trajectory, even when more efficient or rational alternatives may exist.	Negative
Emergency Case	• In case of a climate emergency (e.g., when climate sensitivity is high), CETs could stabilize temperatures • CETs could serve as a back-up plan or insurance against rapid, intense climate impacts • CETs could avert the worst effects of catastrophic climate events	Remark: ambivalent argument, depending if a respondent perceives SAI as a suitable technological fix in case of an emergency. Compared to the “Lesser-evil” this argument is more general and time-pressure is more decisive.	Ambivalent
Lesser-evil	• In a hypothetical scenario there may be a future situation where the deployment of CETs are necessary to prevent catastrophic climate change • In such a scenario the worst impacts of not deploying CETs may be worse than the risks associated with deploying it • CETs would be used as a last resort to avoid the worst impacts of climate change	In contrast to the “Buying Time” argument this argument emphasizes a negative hypothetical scenario and needs to include a comparison. The “Lesser-evil” argument should be often coupled with the “Emergency Case” argument.	Ambivalent

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=5272391

Table 1 (continued).

Buying Time	• CETs could be used as a temporary stopgap to buy time, e.g., for extending climate tipping points • CETs aims to bridge the gap until global mitigation policies become effective • CETs should only be time-limited until its goal is reached and should not lead to decreasing mitigation efforts	Compared to the “Lesser-evil” argument this argument is more general and highlights that CE should only buy time (paralleled by mitigation efforts) and / or should be limited.	Positive
Side-effects not predictable	• Uncertainties in CETs deployment cannot be substantially reduced through research • Deployment of these technologies is considered morally wrong due to these uncertainties • CETs may potentially worsen climate change instead of mitigating it (or increases human health risks)	Respondents could highlight here all kinds of possible side-effects, like increasing lung cancer, because of SAI, but not an unfair distribution of effects (“Unfair distribution of effects and power”).	Negative
Unfair distribution of effects and power	• CETs may disproportionately affect various communities and regions • This can result in unjust distributions of regional climate offsets, costs, and negative side-effects • Areas that have contributed least to climate change may bear most of these technologies' impacts	This argument highlights unfair distribution of effects and not unseen side-effects in general (“Side-effects unseen / not predictable”).	Negative
Hubris Argument	• Not engage in CETs, because the scope of the endeavor is beyond human understanding (virtue perspective) • CETs lack guaranteed effectiveness and full predictability of side effects (consequentialist perspective) • It demonstrates arrogance and self-deceit resulting from an unjustified confidence in knowledge and power beyond what is reasonable for humans	The argument can highlight the hubris for a single human (virtue) or the principle unpredictability of CETs side-effects (consequentialist) and leads to the conclusion not to engage in CETs (different to “Side-effects unseen / not predictable”).	Negative
Betrayal of Divine Creation	• Using CETs is a betrayal of Earth's purpose as given by a higher power (e.g., God). • CETs could signify a move toward "ending nature" and eliminating the world's inherent "wildness" (e.g., pure nature)	The argument highlights compared to the “Hubris Argument” a betrayal of a higher power / entity like God or the purity of nature.	Negative

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=5272391

Table 1 (continued).

Informed Consent	• CETs research and deployment require broad and well-informed consent • Consent should involve representatives of all potentially affected parties (just procedure) • All citizens have a legitimate stake in controlling the "global thermostat"	Remark: ambivalent argument, depending if a respondent perceives that an informed consent of all affected parties is possible. Respondents do not need to emphasize the second or last bullet point.	Ambivalent
Do it Alone	• A determined group of nations can deploy CETs, which benefit the entire world • Long-term cooperation or agreement from all nations may not be necessary	This argument emphasizes that the unilateral use of CETs is positive and not negative ("Risk of Unilateral Use").	Positive
Risk of Unilateral Use	• Research and development of CETs, especially SAI, may result in unilateral deployment with catastrophic consequences • Unilateral climate engineering can lead to political destabilization or be used for hostile purposes • CETs could even independently pursued by wealthy individuals or corporations	Compared to the "Do it Alone" argument this argument is negative and not explicitly highlighting the dual use of the technology as a potential weapon or strategic advantage ("Dual Use").	Negative
Dual Use	• CETs have the potential to modify the weather and therefore could be used as potential weapons • Nations could seek strategic advantage through climate modification methods	Compared to "Risk of Unilateral Use" the argument emphasizes that SAI could be used as a potential weapon or strategic advantage.	Negative
Risk of Governance	• Legal mechanisms for managing CETs, particularly SAI, pose a major challenge • A globally legitimate CETs regime would demand substantial geopolitical stability • SAI technology would need to be safeguarded against involuntary termination (e.g., by terrorist attacks)	The argument is quite broad, highlighting legal issues, geopolitical stability, or possible attacks on the SAI technology. However, when the issue of the long time frame is emphasized the "Long-Term Control" argument should be used.	Negative

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=5272391

Table 1 (continued).

Long-Term Control	• Social systems and institutions may struggle to control CETs over long time scales • Effective management is needed until greenhouse gas emissions are sufficiently reduced and SAI can be withdrawn	Compared to “Risk of Governance” the argument emphasizes the problem of long term control over time. Respondents do not need to emphasize the second bullet point.	Negative
Termination Problem (also “Not Addressing Root Problem”)	• In the absence of effective emissions reduction efforts, greenhouse gasses will continue to accumulate even if temperatures are artificially cooled through SAI • Therefor abrupt termination of SAI may result in rapid, catastrophic climate change, because of large concentration of atmospheric CO2	Remark: a potential termination problem exists only if insufficient mitigation efforts have been made. Therefore, SAI only treats symptoms (rising temperatures), but not causal problems (rising CO2 concentration).	Negative
Greater Good[a]	• If CETs are doing more good than harm then CETs should be deployed. • There could be a "moral obligation" or it could be in general "moral right" to use CETs (deontological perspective). • The technology is for the "greater good" or "maximizes benefits" for society (consequentialist perspective).	Compared to the “Lesser-evil” this argument is rather positive and no negative harms / side-effects are mentioned. There is no comparison (no hypothetical scenario). The argument can highlight the general obligation (deontological) or the positive consequences (consequentialist) using this technology.	Positive

Note. [a]This ethical argument was added after the first coding process (see below). Single ethical arguments like the last two arguments are only specific for SAI and not to “negative emissions” technologies in general.

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=5272391

Importantly, the evaluation of each single CET should consider its placement within a comprehensive climate portfolio, taking into account both the planned scale of its deployment and other climate options such as mitigation or adaptation measures (cf., Aldy et al., 2021; Gardiner, 2010; Neuber, 2018; Ott & Neuber, 2020). To inform such a comprehensive climate portfolio, we propose a study design and appropriate statistical procedures for online-studies in the following sections in order to empirically identify and investigate ethical concerns regarding SAI.

2. Study Design and Statistical Procedures

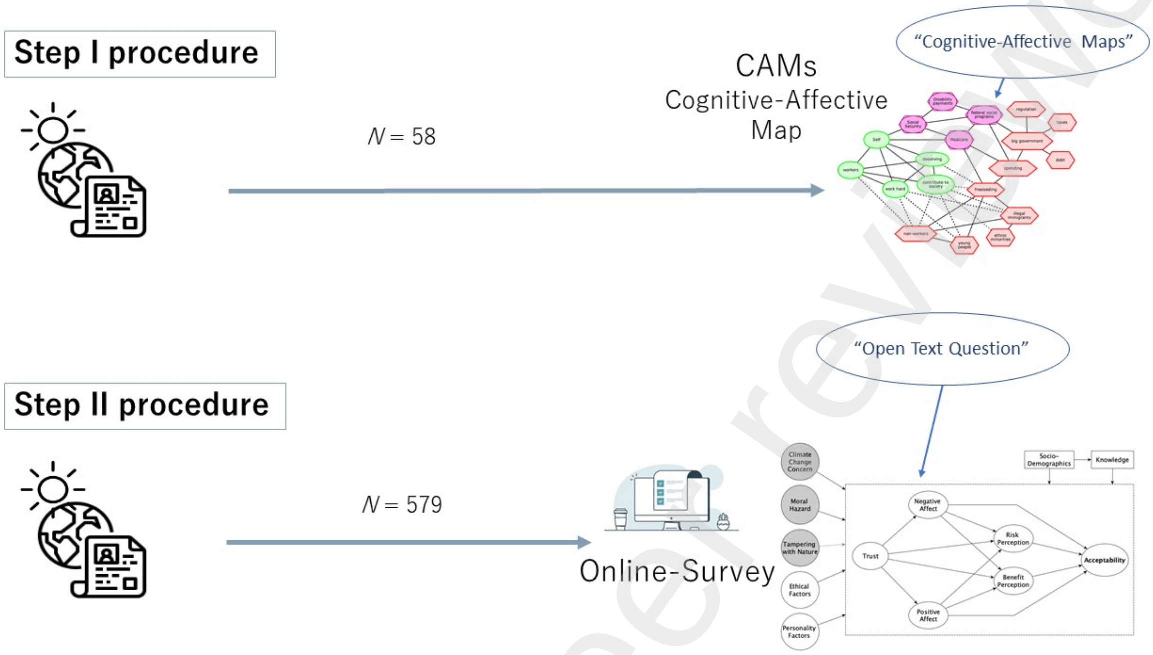
In the current paper, we re-analyse data of a previous study (Fenn et al., 2023), focusing on the identification of ethical concerns of laypersons regarding SAI based on two different types of data (see blue circles within Fig. 1). Participants were informed about the SAI technology by a pre-tested scenario text, describing possible benefits and risks (see https://osf.io/87w6g). The study was composed of two central steps: (a) CAMs were collected at the first measurement time point with a sample size of 58 participants. (b) At the second time point, a large-scale survey with a final sample size of 579 participants was conducted (for details, see Fenn et al., 2023).

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=5272391

Figure 1

Representation of the study design adapted from Fenn, et al. (2023), page 7.



Note. Within the circles, the two types of collected data analyzed in this article ("Cognitive-Affective Maps", "Open Text Question") are highlighted.

This complex study design allows for a multi-method approach to combine heterogeneous sources of data to inform the overall research question (cf., Johnson & Onwuegbuzie, 2004; Steegen et al., 2016). A variety of analytical methods were employed to process and interpret the data, with each method tailored to the specific data type, which are explained in more detail in the respective results sections.

2.1. Scenario-text approach

A balanced and pre-tested scenario text describing the operational principles and different advantages and disadvantages of the SAI technology was created (see Fenn et al., 2023). We considered it necessary to inform the participants about the SAI technology, because multiple articles have reported a relatively low knowledge regarding climate

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=5272391

engineering in general (e.g., Burns et al., 2016; Carlisle et al., 2020; Cummings et al., 2017; Merk et al., 2015). Importantly, we described SAI in the scenario text as imitating nature (e.g., by comparing the effect of SAI to that of volcanoes) to make the scenario text more easily understandable. However, this could have also influenced the perceived naturalness of the technology and could have artificially increased the acceptability (e.g., Corner & Pidgeon, 2015; Thomas et al., 2018). Such an effect is closely linked to the "Natural-is-better" heuristic, whereby nature mostly evokes positive emotions (Siegrist & Árvai, 2020; Siegrist & Hartmann, 2020).

2.2. Cognitive-Affective Maps

CAMs were collected in the Step I procedure (compare Fig. 1) in the study by Fenn et al. (2023). CAMs are a research method encompassing both qualitative and quantitative data-dimensions and can be viewed as a variant of mind maps (Reuter et al., 2022; Thagard, 2010). Participants used our recently developed tools (Fenn et al., under review)[3] to draw their CAM online. An exemplary CAM from the data set is shown in Fig. 3. A CAM consists of concepts and connections, freely drawn by participants to represent their associations. Each concept is assigned an affective connotation by participants on a scale ranging from [-3 to 3], indicating whether the concept evokes positive (green), negative (red), neutral (yellow), or ambivalent (purple) emotions. This visualization provides insights into the emotional valence associated with each concept as perceived by participants. Furthermore, it is possible to write comments to the drawn concepts to further elaborate the drawn concepts. Furthermore, CAMs permit to specify the strength and directionality of connections between these concepts. As such, CAMs can be described as a weighted directional network, which

[3] The interested reader is invited to try out our CAM tools online, as a detailed description of the tools is beyond the scope of this article, see: https://drawyourminds.de/

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=5272391

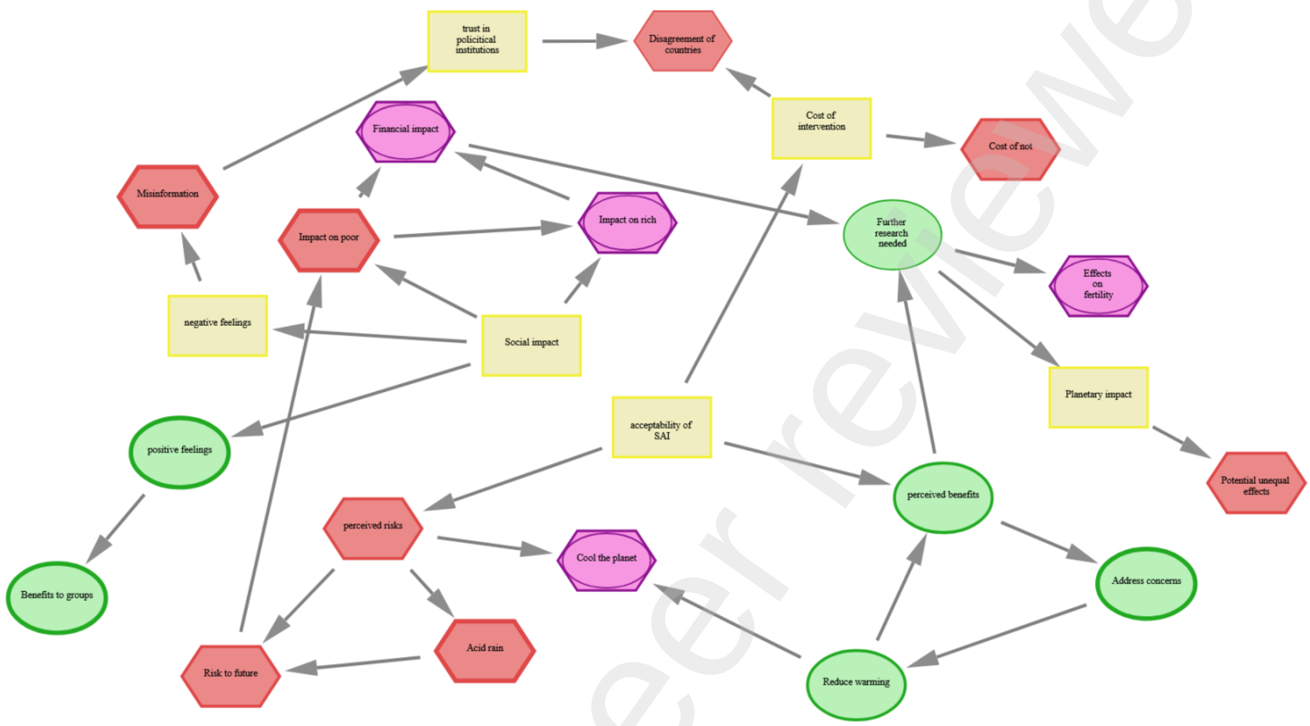
can be analyzed by procedures of network analysis (Bianconi, 2018; Fenn et al., under review; Newman, 2018), whereas the semantic content (written texts, comments) can be analyzed by means of Qualitative Content Analysis (Kuckartz & Rädiker, 2022; Mayring, 2022) and LLMs (Hussain et al., 2024; Tunstall et al., 2022).

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=5272391

Figure 3

Exemplary CAM with an average valence of -0.16 drawn by a participant.



Note. In this CAM, different concepts already indicate ethical arguments, which can be found in Tab. 1. For example, “Disagreement of countries” corresponds to the ethical argument of “Risk of Governance”.

2.3. Open-Text

In the Step II procedure in the study by Fenn et al. (2023), participants first read the scenario text and then immediately answered the following question: “When, in your opinion, is the described ‘Stratospheric Aerosol Injection’ technology morally right?”. Additionally, participants were provided a general definition of morality (based on Jacobs, 2002; Pieper, 2017). Participants were forced to take at least one minute to answer this open text question. These open text answers were analyzed according to the procedure of qualitative content analysis. To support the raters to code the text material applying qualitative content analysis,

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=5272391

a YouTube video with coding instructions was created[4]. Qualitative content analysis is a systematic method of analyzing text data that involves coding and categorizing the content to derive themes and patterns. It emphasizes a structured approach that includes several stages, such as preparation, forming categories, and coding the material (Kuckartz & Rädiker, 2022; Mayring, 2022). To code the text material, coding guidelines which contain all the theoretically derived ethical arguments (see Tab. 1) were developed. Then, a three step coding procedure followed (see "Results" section for details).

3. Results

The following statistical analyses were analyzed using R (R Core Team, 2020), Mplus (Muthén & Muthen, 2017) and Python (Van Rossum & Drake, 2009). The analysis scripts, which have been written in the form of text annotated reproducible scripts by using the "rmarkdown" package in R (Xie et al., 2018) and the Jupyter notebook for Python (Kluyver et al., 2016), are publicly accessible via the Open Science Framework (OSF) at https://doi.org/10.17605/OSF.IO/ANXG6.

The subsequent sections employ a deductive-driven qualitative content analysis to summarize insights from the CAMs and open-text responses. In contrast, the final results section adopts a predominantly inductive, data-driven approach, leveraging LLMs to synthesize the most prevalent ethical arguments and underlying patterns of reasoning. Key findings are summarized in Table 2 presented in the "Discussion" section.

3.1. Cognitive-Affective Maps

58 participants (mean age 38, *SD* = 10.40, 47% female) drew the CAMs online by using recently developed tools (Fenn et al., 2024). We pre-defined the concept "acceptability

[4] see YouTube Video: https://www.youtube.com/watch?v=725flcytGJw

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=5272391

of SAI" in the center of the CAM. On the top, five additional concepts ("positive feelings", "negative feelings", "trust in political institutions", "perceived risks", and "perceived benefits") were presented. Participants were able to move or delete the five additional pre-defined concepts and were technically required to draw at least 24 concepts in total (see in detail Fenn et al., 2023). The dataset consists of 58 CAMs, where participants drew an average of 25.4 concepts each, with 34% positive, 46% negative, 12% neutral, and 8% ambivalent. On average, 44.21 connectors were drawn per CAM. The mean valence across all CAMs was -0.33, and at least one predefined concept was removed in 14% of the CAMs. Detailed descriptive statistics are provided in Appendix A.

3.1.1. Data preparation

While summarizing the CAM data for this article, we focused on ethical arguments within the individually drawn CAMs. Thereby, each drawn concept can be assigned to one of the 21 different ethical arguments (see Tab. 1). The CAM data were summarized by our developed Data Analysis Tool, and we iteratively summarized the semantic content of the CAMs (Fenn et al., 2023, under review). The initial 1,473 concepts (consisting of 1,063 unique concepts) were condensed to a final set of 41 concepts.

3.1.2. Data analysis

For this analysis, we focus only on ethical arguments (see Tab. 1) within the CAMs. 19 of the 21 ethical arguments were mentioned at least by one participant – only two ethical arguments "Emergency Case" and "Unstoppable Deployment if researched" were not mentioned. Overall 724 of the total 1,473 drawn concepts could be assigned to ethical arguments, whereby the argument "Side-effects not predictable" was mentioned most frequently (264 times), followed by the positive ethical argument "Greater Good" (143). The ethical argument of "Side-effects not predictable" encompasses all kinds of negative

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=5272391

evaluations, e.g., SAI could cause “acid rain”, increase “health risks”, and “side-effects [are] not fully resolved”. In contrast, “Greater Good” highlights that SAI could have “environmental benefits” or could even “save the world”. Participants also frequently highlighted the “Risk of Governance” (e.g., unclear who is “accountable” or risks of “political instability”), which is linked to the negatively perceived argument of “Informed Consent” highlighting the lack of “political consensus” and the impossibility “to get everyone to agree on [SAI]”. The argument of “Long-Term Control” is also negatively perceived, conveying that it “could be hard to get all countries to commit for a long period of time”. However, there is a mixed perception regarding SAI as a “Technological Fix”. Some participants viewed it as “a solution where we don’t have to make changes to our everyday life,” while more frequently, participants highlighted concerns that SAI could foster a “Moral Hazard”. A table with the frequencies of all the mentioned ethical arguments and a few examples from participants can be found in Appendix B, a complete wordlist can be found on OSF (https://osf.io/rhxfn).

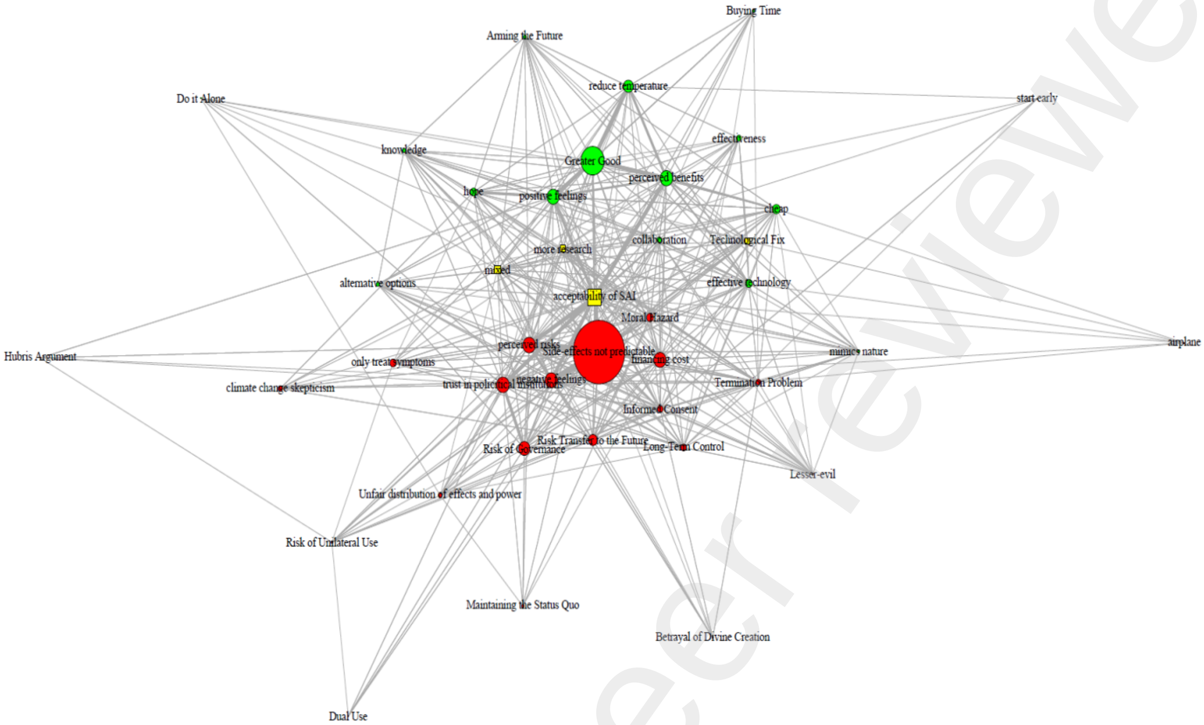
All these ethical arguments are interrelated, which can be seen in Fig. 4, where we show an aggregated network based on all 58 CAMs. The relative size of the concepts and the thickness of the connections indicate the frequency of the drawn concepts and the frequency of the pairwise connections, respectively. On average, participants mentioned 12.48 (SD = 3.40) ethical arguments in their CAMs.

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=5272391

Figure 4

Aggregated CAM consisting of N = 58 CAMs.



Note. A zoomable PDF file can be found on OSF (https://osf.io/xr62c). The color is indicative of the average valence of the concepts, whereby yellow represents neutral, green positive, and red negative concepts. If the average valence of a concept is within [-0.5, 0.5] a concept was drawn as neutral. Remark: Figure was adjusted from Fig. 3 in Fenn et al. (2023) on page 11.

3.1.3. Discussion of CAM results

Most frequently, participants highlighted negative ethical arguments, especially all kinds of possible side-effects. As visible in Fig. 4, these ethical arguments are strongly interrelated, whereby participants emphasize the importance of governance related ethical arguments (“Risk of Governance”, “Informed Consent”, “Long-Term Control”, “Termination Problem”). Such concerns could inform possible issues of governing solar geoengineering (e.g., Flegal et al., 2019; MacMartin et al., 2019). As can be seen in the upper part of Fig. 4, participants drew multiple positively assessed concepts regarding SAI (e.g., “Greater Good”).

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=5272391

These concepts are connected to other summarized concepts, like SAI being capable of reducing the temperature or being relatively cheap. About 9% of participants mentioned "Betrayal of Divine Creation" as strongly negative, arguing that SAI is not acceptable because SAI is like "playing God" or violates the purity of nature. Such an argument could be a strong moral heuristic when making (ethical) decisions (cf., Schwartz, 2016). Also, 16% of the participants drew positive concepts that SAI is mimicking nature, e.g., that SAI "would have similar effects on the atmosphere as volcanoes" (see Appendix B). Such a finding emphasizes the importance of how CETs are framed in general (see section 2.1.).

3.2. Open Text

In total, we had a final sample size of 579 participants (*M* = 40 years, *SD* = 13.26, 47% female). Three participants provided no answer and two participants indicated that they did not know how to answer the open text question regarding the morality of SAI. Removing participants with no answers we have 576 answers varying in length from 1 to 171 words (mean number of words: 36.26, *SD* = 21.96). By applying the Python module VADER (for Valence Aware Dictionary for sEntiment Reasoning; Hutto & Gilbert, 2014), we computed sentiment scores which refer to the emotional tone expressed in the text answers. In total, 273 answers were negative, 246 positive and 57 neutral. Descriptively, neutral arguments seemed less elaborated and had on average only 17.79 words (*SD* = 11.35). The following section outlines the three-step procedure used to categorize ethical arguments within the text responses.

3.2.1. Data preparation

The summary of the open text data was based on qualitative content analysis (Mayring, 2022) using the open access QCAmap application (Fenzl & Mayring, 2017). To summarize the data, we followed the strict step model of the procedure of "deductive

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=5272391

category assignment"[5]. Based on existing theories, a category system was defined (see Tab. 1), followed by seven raters coding 5% of the text answers in a first step. This procedure led to minor adjustments of the category system. Importantly, a new ethical argument ("Greater Good") was added to code text answers emphasizing that the technology is doing more good than harm without mentioning negative side-effects (for details see: https://osf.io/evzwm). Finally, in the last step, the complete text data was coded by seven raters, whereby six of these raters also participated in the first coding process.

Multiple quality criteria were applied to check the quality of the summarizing process of the text data (motivated by Kuckartz & Rädiker, 2022; Mayring, 2022). The content validity (Moosbrugger & Kelava, 2020) of the category system was reflected within team discussions involving two ethics experts to ensure that the ethical arguments (categories) were constructed in such a way that they capture central reasoning discussed in the philosophy of ethics. For the first coding process, we tested for the inter-rater reliability, computing Fleiss' kappa (Fleiss et al., 2013) to check for discrepancies between the seven raters for every single open text answer coded. On average, the reliability was substantial with $\varkappa = .75$ ($p < .01$). To improve the coding of the complete text data (final step), we followed the procedure of "subjective assessment" (Guest et al., 2012), whereby discrepancies were discussed in a group discussion with all the raters until consensus was reached. The category system was adjusted respectively.

3.2.2. Data analysis

In the following, only 553 of 576 text answers are considered, because in 15 (2.6%) text answers no ethical argument was assigned and in 11 (1.9%) no consensus was found regarding the coding of the respecting text answer. Motivated by Pokorny et al. (2018), we

[5] Accessible in the QCAmap application, see: https://www.qcamap.org/ui/assets/tutorials/en/Steps_Rules_Deductive.pdf; last accessed on January 9, 2025

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=5272391

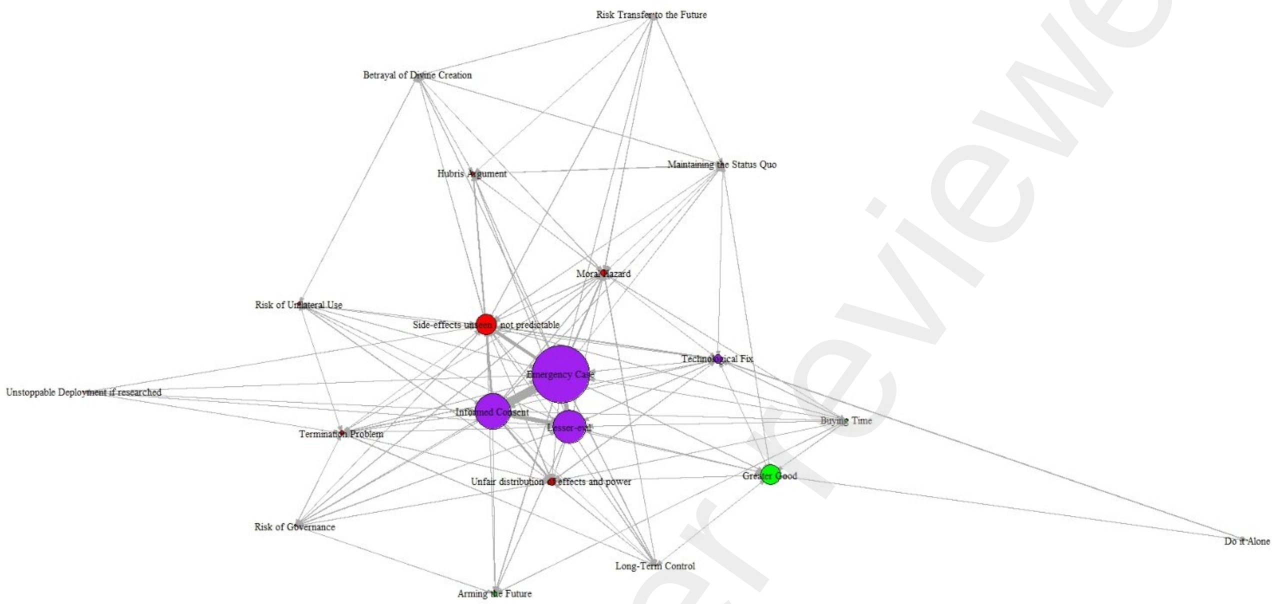
visualized the coding of the 553 text answers in the form of a network in Fig. 5. The three most frequently mentioned ethical arguments were "Emergency Case" (assigned 202 times), followed by "Informed Consent" (126) and "Lesser-Evil" (119). These three ambivalent ethical arguments (see Tab. 1) were often mentioned within the same text answers, linking them argumentatively. For example, a participant linked the ethical arguments of "Emergency Case" and "Lesser-Evil" by stating if "climate change has reached a point where human life is affected to such a degree that large numbers of people are facing hardships [..] [SAI] would be a last resort" (quote of one participant). Using this technology in "an imminent catastrophic climate emergency" would only be morally right if this action is "agreed on by all the major or world leading countries of the world" (quote by another participant) highlights a connection between "Lesser-Evil" and "Informed Consent". Thereby, open text answers varied regarding who needs to agree (e.g., global, all affected, or leading countries; see word clouds of given answers: https://osf.io/5ztcj). Interestingly, participants also frequently mentioned the "Greater Good" of the technology, expressing a generally positive attitude towards the technology, e.g., "it will benefit society, and help in general" (quote of another participant). In addition, the argument of the general unpredictability of the technology (ethical argument of "Side-effects not predictable") was mentioned as often as the "Greater Good" argument. The ethical argument of "Dual Use" has not been mentioned by a single participant and some ethical arguments were mentioned less than ten times (for details see table in Appendix C).

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=5272391

Figure 5

Covariation of ethical arguments within open text.



Note. A zoomable PDF file can be found on OSF (https://osf.io/mxt89). The color coding represents the "General Evaluation" column in Tab. 1 with green = positive, red = negative, yellow = neutral, and purple = ambivalent. The frequency of drawn concepts and the number of pairwise connections is proportional to the size of the concept and the thickness of the connections, respectively.

3.2.3. Discussion of Open Text results

Participants mentioned ambivalent ethical arguments most frequently in their open text answers, frequently highlighting that SAI would be the "Lesser-Evil" if there would be an "Emergency Case". At the same time, it was also emphasized that the informed consent of all, or at least of the majority of countries, would be necessary. Unlike the CAM data, participants highlighted possible negative-side effects and general benefits of the SAI technology ("Greater Good") to an equal extent. Also, governance-related issues were hardly mentioned. Thus, ambivalent ethical arguments strongly dominated the text answers.

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=5272391

Importantly, multiple participants already emphasized that the ethical arguments of “Emergency Case” and “Lesser-Evil” are strongly related (e.g., Gardiner, 2010b; Ott & Neuber, 2020), whereby the lesser evil argument has the potential to become a self-fulfilling prophecy, as preparing for a horrific scenario may inadvertently lead to its occurrence.

3.3. Testing Large Language Models for Summarizing Ethical Arguments

To analyze coding assignments derived from the CAM alongside open-text data, we utilized a LLM, specifically "Llama-3.1-70B-Instruct" (Dubey et al., 2024) , in two distinct applications: First, the model was tasked with synthesizing marked text passages - segments of text identified as relevant during the coding process - to generate comprehensive summaries of the corresponding ethical arguments. This initial application facilitated the distillation of extensive textual datasets into concise, interpretable summaries tailored to the ethical coding guidelines under examination. Subsequently, the LLM was applied again to identify commonalities and differences by composing a synthesized paragraph summarizing the shared themes and discrepancies between our formal definition of the ethical arguments (see Table 1) and laypersons' associations with the respective ethical argument. Both times, we set up well-structured prompts by including contextual data, explanations of the data structure, and including the respective definition of the ethical argument into the prompts, adhering to established best practices in the literature (Dai et al., 2023; Liu et al., 2023; White et al., 2023). A non-technical introduction to LLMs is provided in Appendix D.

3.3.1. LLM-Generated Summaries of Ethical Arguments

Table 2 presents the outcomes of the initial prompting of the LLM, offering a structured and concise synthesis of two exemplary ethical arguments: *Moral Hazard* and *Technological Fix* (for all other ethical arguments see table on OSF: https://osf.io/jx4hc**)**. For each of the two

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=5272391

datasets – Open Text and CAM data – the LLM-generated summaries encapsulate the most salient themes and patterns identified within the diverse associations articulated by laypersons in response to these ethical arguments.

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=5272391

Table 2

Laypersons' Associations with Ethical Arguments on SAI, Synthesized Using LLM

Ethical Argument	Summary Open Text	Summary CAM
Moral Hazard (also: “Undermining Better Options”)	• CE might reduce enthusiasm for mitigation efforts, as it may be viewed as a last resort for continued pollution. • Focusing on CE research may divert resources from essential emission abatement and adaptation strategies. • CE does not address climate change's root cause (CO2 emissions) and may lead to complacency. • Tackling excessive CO2 emissions through reduced energy use, green energy, and lifestyle changes is crucial. • Prioritizing CE over mitigation efforts is morally problematic, as it doesn’t resolve pollution or high CO2 levels.	• CE could lead to complacency, as people may view it as a solution and reduce efforts to address climate change. • It may be seen as a distraction from real issues, diverting resources away from long-term sustainable change. • Governments may choose CE for its simplicity, reducing investment in sustainable energy and climate action. • CE doesn't encourage innovation and may lead to political lethargy, preventing necessary action against climate change.
Technological Fix	• Technological fixes like SAI are justified only after all other options have been explored, as they treat symptoms, not the root cause. • Without addressing CO2 emissions and behavioral changes, such fixes are counterproductive and not sustainable. • Reducing CO2 emissions and promoting behavioral changes should be prioritized over temporary technological fixes. • Technological fixes can lead to complacency, reducing motivation for necessary systemic changes. • Such solutions should be last resorts, used only when efforts to reduce emissions fail.	• Technological fixes like SAI are seen as quick solutions but may delay necessary lifestyle and behavioral changes. • Adoption may be driven by political indecision or lack of will to implement more comprehensive climate solutions. • While urgent, technological fixes don’t address the root cause of climate change and may offer only temporary relief. • Their low cost and simplicity may influence decisions, but they are not always the most effective long-term solution. • They may reduce personal responsibility, offering an easy way out instead of encouraging behavioral change.

Note. All other layperson’s associations to the ethical arguments can be found on OSF, see: https://osf.io/jx4hc

3.3.2. Exploring Ethical Parallels: Formal Definitions vs. Lay Perspectives

The outcome of this LLM prompt is a synthesized and concise summary that systematically identifies and highlights the shared themes and significant discrepancies between laypersons’ interpretations and our formal definitions of the ethical arguments (see table available on OSF: https://osf.io/sncjh). This comparative synthesis highlights notable differences between our theory-driven definitions of the ethical arguments (see Table 1) and laypersons’

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=5272391

associations with these arguments: These differences arise from contrasting emphases on moral, social, and technical dimensions. In our definitions we predominantly focused on structured conceptual frameworks, such as resource diversion (Moral Hazard), future dilemmas (Risk Transfer), and the theoretical last-resort nature of climate engineering (Lesser Evil). Conversely, laypersons emphasize immediate moral implications, societal trust, and tangible impacts. Lay perspectives especially highlight issues such as climate engineering's failure to address root causes, political inertia, global conflicts, and inter-nation mistrust, alongside the need for systemic change and accountability. Context-specific concerns - such as temperature thresholds (Emergency Case) and catastrophic consequences (Risk Transfer) - also feature more prominently in laypersons' interpretations. Overall laypersons' arguments are framed through lived experiences and practical concerns, often introducing notions of fairness, equity, and societal trust.

3.3.3. Discussion of LLM results

A comparison of the LLM-generated summaries in Section 3.3.1 reveals both similarities and differences in the emphases of CAM and open-text data. Both datasets consistently highlight the moral and ethical implications of CE sharing concerns, for example, about its potential to divert attention from addressing root causes and undermining sustainable climate solutions. However, CAM data more frequently emphasizes geopolitical and social dimensions, such as global conflicts and inter-nation mistrust, whereas open-text data focuses on individual responsibility and the moral necessity of addressing systemic issues like resource diversion and behavioral change. The greater emphasis on geopolitical dimensions in CAM data could be attributed to the predefined inclusion of the concept "trust in political institutions", which shaped the framing of responses in that dataset. As shown in Section 3.3.2., notable

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=5272391

discrepancies exist between laypersons' interpretations and our formal definitions of ethical arguments. These differences reflect contrasting emphases: while formal definitions focus on structured, theoretical constructs such as resource diversion and risk transfer, laypersons prioritize immediate moral implications, lived experiences, and tangible outcomes. These results underscore the unique contribution of laypersons' associations in expanding the discourse around ethical arguments.

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=5272391

4. Conclusion

In this article, we demonstrate the value of integrating two distinct data types – open text responses and CAMs – to explore laypersons' ethical concerns regarding the use of SAI. CAMs offer a structured visualization of ethical concerns, identifying a broad spectrum of issues ranging from "trust in political institutions" to "mimicking nature". Thereby participants structure ethical arguments in the process of drawing CAM, which is related to the theoretical concept of ethical coherence (Thagard, 1998, 2000). In contrast, open-text data revealed mainly ambivalent arguments (e.g., "Emergency Case," "Lesser-Evil," "Informed Consent"). Open text and CAMs could enable future researchers to identify central ethical arguments or even *master-narratives* regarding (such) emerging technologies, such as the notion that deploying these technologies is like "Opening Pandora's box" (Davies & Macnaghten, 2010; Macnaghten et al., 2019). It could be the case that CAMs foster deliberative thinking and enable participants to structure complex ethical arguments in the form of complex interconnected maps (see Vink et al., 2016). Methodological differences are summarized alongside our key findings in Table 3.

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=5272391

Table 3

Overview of the type of data, the main outcomes and reflection of the results of the two types of data

	Cognitive-Affective Maps	Open Text
Type of Data	qualitative and quantitative	qualitative
Main Outcomes	* broad range of ethical arguments identified, including governance related arguments * ethical arguments are linked to other predefined concepts (e.g., "trust in political institutions") * arguments like "feeling that SAI mimics nature" or "brings hope" could influence the ethical argumentation	* mainly ambivalent ethical arguments identified * three ambivalent ethical arguments ("Emergency Case", "Lesser-Evil", "Informed Consent") are argumentatively interlinked in the text answers
Reflection	* depending on the pre-defined concepts, participants probably highlight different ethical arguments * participants were required to draw 24 concepts, potentially leading to the high number of possible negative side-effects that were mentioned	* by answering a general question regarding the morality of SAI, participants might be inclined to think about the argument that SAI could be used in case of an emergency * ethical arguments, like governance issues, do not seem to be mentally present

Note. The "Type of Data" can be "qualitative" (e.g., answer to open-ended survey questions resulting in text) or "quantitative" (e.g., drawing a network resulting in specific network parameters).

Referring to the detailed review by Reynolds & Horton (2020) the findings outlined in this article yield insights for the analytical problems of the Earth System Governance framework (Biermann et al., 2010; Burch et al., 2019): Laypersons in the CAM data highlighted governance related ethical arguments and emphasized central problems like inter-nation mistrust, unclear accountability, lack of consensus or potential political instability highlighting problems of a potential future governance architecture. Further the open-text data revealed concerns about equitable participation ("informed consent of all countries") and thereby pointing to the moral legitimacy of decision-making, emphasizing concerns of potential power asymmetries. The frequent "Lesser-Evil" and "Emergency Case" arguments underscore distributive and intergenerational justice challenges, while the LLM-mediated

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=5272391

synthesis emphasizes how lay perspectives foreground practical burdens and potential benefits for vulnerable populations. Lastly laypersons imagined future scenarios, whether hopeful ("Greater Good") or alarming ("Betrayal of Divine Creation"), illustrating the power of narrative imaginations.

4.1. Future research and limitations

In our opinion, the ethical concerns of emerging CETs should be assessed continuously, because if an ethical assessment would wait until sufficient information regarding the (side-)effects of a CET becomes available, a technology would be deeply ingrained in society and the potential for making revisions would be strongly limited (called the "Collingridge dilemma", Collingridge, 1980; Möller & Grießhammer, 2022). This justifies an early ethical assessment of such technologies, even if knowledge of laypersons regarding CETs is low (Grunwald & Hillerbrand, 2021; Palm & Hansson, 2006). Such a perspective emphasizes the need to conduct future studies, e.g., ideally as a "tracker technology assessment" (see Böschen et al., 2021; Lucivero, 2016) to inform the empirical ethical assessment of CETs during different development stages of these technologies. For example, "Moral Hazard" is of particular importance at an early research stage while the "Termination Problem" is particularly important when such a technology would be finally implemented (see Preston, 2013). Future research should systematically examine whether increasing participants' knowledge about CETs influences their ethical concerns and their envisioning of potential futures involving such technologies. Here one might conjecture, for example, information-choice questionnaires (e.g., Gregory et al., 2016; ter Mors et al., 2013). In addition, a future study could provide a more straightforward comparison of CAMs and text data. In the present study, such a comparison was limited due to fundamental differences in methodological design - particularly the pre-defined concepts in the CAM approach, which

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=5272391

may trigger different cognitive associations and thereby shape the ethical concerns participants express (see Lichtenstein & Slovic, 2006). If CAMs provide similar information as text data, we would recommend applying CAMs with different sets of pre-defined concepts because such data can be semi-automatically summarized.

To inform a comprehensive climate portfolio, the ethical concerns of all important CETs should be assessed by multiple stakeholder groups (cf., Aldy et al., 2021; Gardiner, 2010; Neuber, 2018; Ott & Neuber, 2020). We therefore encourage future researchers to adopt the methodology proposed in this study, along with the accompanying online resources, to advance empirically informed ethics of CETs. Incorporating laypersons' perspectives can enhance the inclusivity and societal relevance of discourse on these technologies, thereby supporting climate policy and anticipatory governance.

Finally we want to stress that research in this domain is particularly needed, as participants in our study referred to the potential deployment of SAI in the context of a future climate emergency, framing it as a "lesser evil" in a hypothetical but severe crisis scenario. This justification for the ultimate use of such a risky technology (e.g., Sovacool et al., 2022, 2023) underscores the critical importance of preemptively avoiding such "emergency situations" so that there is still room for ethical discussions to govern such technologies before potentially irreversible measures become necessary (cf., Gardiner, 2011; Ott, 2011).

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=5272391

References

Aldy, J. E., Felgenhauer, T., Pizer, W. A., Tavoni, M., Belaia, M., Borsuk, M. E., Ghosh, A., Heutel, G., Heyen, D., Horton, J., Keith, D., Merk, C., Moreno-Cruz, J., Reynolds, J. L., Ricke, K., Rickels, W., Shayegh, S., Smith, W., Tilmes, S., … Wiener, J. B. (2021). Social science research to inform solar geoengineering. *Science*, *374*(6569), 815–818. https://doi.org/10.1126/science.abj6517

Anderson, K., & Peters, G. (2016). The trouble with negative emissions. *Science*, *354*(6309), 182–183. https://doi.org/10.1126/science.aah4567

Barrett, S. (2008). The Incredible Economics of Geoengineering. *Environmental and Resource Economics*, *39*(1), 45–54. https://doi.org/10.1007/s10640-007-9174-8

Betz, G., & Cacean, S. (2012). *Ethical Aspects of Climate Engineering*. KIT Scientific Publishing.

Bianconi, G. (2018). *Multilayer Networks: Structure and Function*. Oxford University Press.

Biermann, F., Betsill, M. M., Gupta, J., Kanie, N., Lebel, L., Liverman, D., Schroeder, H., Siebenhüner, B., & Zondervan, R. (2010). Earth system governance: A research framework. *International Environmental Agreements: Politics, Law and Economics*, *10*(4), 277–298. https://doi.org/10.1007/s10784-010-9137-3

Biermann, F., Oomen, J., Gupta, A., Ali, S. H., Conca, K., Hajer, M. A., Kashwan, P., Kotzé, L. J., Leach, M., Messner, D., Okereke, C., Persson, Å., Potočnik, J., Schlosberg, D., Scobie, M., & VanDeveer, S. D. (2022). Solar geoengineering: The case for an international non-use agreement. *WIREs Climate Change*, *13*(3), e754. https://doi.org/10.1002/wcc.754

Bindé, J. (2000). Toward an Ethics of the Future. *Public Culture*, *12*(1), 51–72. https://doi.org/10.1215/08992363-12-1-51

Böschen, S., Grunwald, A., Krings, B.-J., & Rösch, C. (2021). *Technikfolgenabschätzung:*

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=5272391

Handbuch für Wissenschaft und Praxis. Nomos Verlag.

Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., Agarwal, S., Herbert-Voss, A., Krueger, G., Henighan, T., Child, R., Ramesh, A., Ziegler, D. M., Wu, J., Winter, C., … Amodei, D. (2020). *Language Models are Few-Shot Learners* (No. arXiv:2005.14165). arXiv. https://doi.org/10.48550/arXiv.2005.14165

Buckley, J. A., Thompson, P. B., & Whyte, K. P. (2017). Collingridge's dilemma and the early ethical assessment of emerging technology: The case of nanotechnology enabled biosensors. *Technology in Society*, *48*, 54–63. https://doi.org/10.1016/j.techsoc.2016.12.003

Burch, S., Gupta, A., Inoue, C. Y. A., Kalfagianni, A., Persson, Å., Gerlak, A. K., Ishii, A., Patterson, J., Pickering, J., Scobie, M., Van der Heijden, J., Vervoort, J., Adler, C., Bloomfield, M., Djalante, R., Dryzek, J., Galaz, V., Gordon, C., Harmon, R., … Zondervan, R. (2019). New directions in earth system governance research. *Earth System Governance*, *1*, 100006. https://doi.org/10.1016/j.esg.2019.100006

Burns, E. T., Flegal, J. A., Keith, D. W., Mahajan, A., Tingley, D., & Wagner, G. (2016). What do people think when they think about solar geoengineering? A review of empirical social science literature, and prospects for future research. *Earth's Future*, *4*(11), 536–542. https://doi.org/10.1002/2016EF000461

Caelen, O., & Blete, M.-A. (2023). *Developing Apps with Gpt-4 and Chatgpt: Build Intelligent Chatbots, Content Generators, and More* (1st ed.). O'Reilly Media.

Callies, D. E. (2019). The Slippery Slope Argument against Geoengineering Research. *Journal of Applied Philosophy*, *36*(4), 675–687. https://doi.org/10.1111/japp.12345

Carlisle, D. P., Feetham, P. M., Wright, M. J., & Teagle, D. A. H. (2020). The public remain uninformed and wary of climate engineering. *Climatic Change*, *160*(2), 303–322.

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=5272391

https://doi.org/10.1007/s10584-020-02706-5

Carr, W. A., & Yung, L. (2018). Perceptions of climate engineering in the South Pacific, Sub-Saharan Africa, and North American Arctic. *Climatic Change*, *147*(1), 119–132. https://doi.org/10.1007/s10584-018-2138-x

Caviezel, C., & Revermann, C. (2014). *Climate Engineering: Kann und soll man die Erderwärmung technisch eindämmen?* edition sigma. https://publikationen.bibliothek.kit.edu/140100230

Chiang, W.-L., Zheng, L., Sheng, Y., Angelopoulos, A. N., Li, T., Li, D., Zhang, H., Zhu, B., Jordan, M., Gonzalez, J. E., & Stoica, I. (2024). *Chatbot Arena: An Open Platform for Evaluating LLMs by Human Preference* (No. arXiv:2403.04132). arXiv. https://doi.org/10.48550/arXiv.2403.04132

Collingridge, D. (1980). *Social Control of Technology*. Milton Keynes: Open University Press.

Corner, A., Parkhill, K. A., & Pidgeon, N. (2011). *"Experiment Earth?" Reflections on a public dialogue on geoengineering: Reflections on a public dialogue on geoengineering* [Working Paper]. Cardiff University. https://eprints.whiterose.ac.uk/82861/

Corner, A., Parkhill, K., Pidgeon, N., & Vaughan, N. E. (2013). Messing with nature? Exploring public perceptions of geoengineering in the UK. *Global Environmental Change*, *23*(5), 938–947. https://doi.org/10.1016/j.gloenvcha.2013.06.002

Corner, A., & Pidgeon, N. (2014). Geoengineering, climate change scepticism and the 'moral hazard' argument: An experimental study of UK public perceptions. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *372*(2031), 20140063. https://doi.org/10.1098/rsta.2014.0063

Corner, A., & Pidgeon, N. (2015). Like artificial trees? The effect of framing by natural

Preprint not peer reviewed

analogy on public perceptions of geoengineering. *Climatic Change*, *130*(3), 425–438. https://doi.org/10.1007/s10584-014-1148-6

Cotton, M. (2014). *Ethics and Technology Assessment: A Participatory Approach*. Springer.

Crutzen, P. J. (2006). Albedo Enhancement by Stratospheric Sulfur Injections: A Contribution to Resolve a Policy Dilemma? *Climatic Change*, *77*(3–4), 211. https://doi.org/10.1007/s10584-006-9101-y

Cummings, C. L., Lin, S. H., & Trump, B. D. (2017). Public perceptions of climate geoengineering: A systematic review of the literature. *Climate Research*, *73*(3), 247–264. https://doi.org/10.3354/cr01475

Dai, S.-C., Xiong, A., & Ku, L.-W. (2023). *LLM-in-the-loop: Leveraging Large Language Model for Thematic Analysis* (No. arXiv:2310.15100). arXiv. https://doi.org/10.48550/arXiv.2310.15100

Davies, S. R., & Macnaghten, P. (2010). Narratives of Mastery and Resistance: Lay Ethics of Nanotechnology. *NanoEthics*, *4*(2), 141–151. https://doi.org/10.1007/s11569-010-0096-5

De Bruin, K. C., Dellink, R. B., & Tol, R. S. J. (2009). AD-DICE: An implementation of adaptation in the DICE model. *Climatic Change*, *95*(1–2), 63–81. https://doi.org/10.1007/s10584-008-9535-5

Debelak, R., Koch, T., Aßenmacher, M., & Stachl, C. (2024). *From Embeddings to Explainability: A Tutorial on Transformer-Based Text Analysis for Social and Behavioral Scientists*. OSF. https://doi.org/10.31234/osf.io/bc56a

Dowling, A. (2018). *Greenhouse Gas Removal* [Monograph]. Royal Society. https://royalsociety.org/topics-policy/projects/greenhouse-gas-removal/

Dubey, A., Jauhri, A., Pandey, A., Kadian, A., Al-Dahle, A., Letman, A., Mathur, A., Schelten, A., Yang, A., Fan, A., Goyal, A., Hartshorn, A., Yang, A., Mitra, A.,

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=5272391

Sravankumar, A., Korenev, A., Hinsvark, A., Rao, A., Zhang, A., … Zhao, Z. (2024). *The Llama 3 Herd of Models* (No. arXiv:2407.21783). arXiv. https://doi.org/10.48550/arXiv.2407.21783

Fenn, J., Gouret, F., Gorki, M., Reuter, L., Gros, W., Hüttner, P., & Kiesel, A. (under review). *Cognitive-Affective Maps extended logic: Proposing Tools to Collect and Analyze Attitudes and Belief Systems*.

Fenn, J., Helm, J. F., Höfele, P., Kulbe, L., Ernst, A., & Kiesel, A. (2023). Identifying key-psychological factors influencing the acceptance of yet emerging technologies–A multi-method-approach to inform climate policy. *PLOS Climate*, *2*(6), 1–25. https://doi.org/10.1371/journal.pclm.0000207

Fenzl, T., & Mayring, P. (2017). *QCAmap: Eine interaktive Webapplikation für Qualitative Inhaltsanalyse*. https://doi.org/10.23668/psycharchives.11259

Fiorino, D. J. (1990). Citizen Participation and Environmental Risk: A Survey of Institutional Mechanisms. *Science, Technology, & Human Values*, *15*(2), 226–243. https://doi.org/10.1177/016224399001500204

Flegal, J. A., Hubert, A.-M., Morrow, D. R., & Moreno-Cruz, J. B. (2019). Solar Geoengineering: Social Science, Legal, Ethical, and Economic Frameworks. *Annual Review of Environment and Resources*, *44*(1), 399–423. https://doi.org/10.1146/annurev-environ-102017-030032

Fleiss, J. L., Levin, B., & Paik, M. C. (2013). *Statistical Methods for Rates and Proportions*. John Wiley & Sons.

Frumhoff, P. C., & Stephens, J. C. (2018). Towards legitimacy of the solar geoengineering research enterprise. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *376*(2119), 20160459. https://doi.org/10.1098/rsta.2016.0459

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=5272391

Funtowicz, S. O., & Ravetz, J. R. (2018). Post-normal science. In *Companion to Environmental Studies*. Routledge.

Gardiner, S. M. (2010a). Ethics and climate change: An introduction. *WIREs Climate Change*, *1*(1), 54–66. https://doi.org/10.1002/wcc.16

Gardiner, S. M. (2010b). Is “Arming the Future” with Geoengineering Really the Lesser Evil? Some Doubts about the Ethics of Intentionally Manipulating the Climate System. In *Climate Ethics Essential Readings* (pp. 284–312).

Gardiner, S. M. (2011). *A Perfect Moral Storm: The Ethical Tragedy of Climate Change*. Oxford University Press.

Gardiner, S. M., & Fragnière, A. (2018). The Tollgate Principles for the Governance of Geoengineering: Moving Beyond the Oxford Principles to an Ethically More Robust Approach. *Ethics, Policy & Environment*, *21*(2), 143–174. https://doi.org/10.1080/21550085.2018.1509472

Gregory, R., Satterfield, T., & Hasell, A. (2016). Using decision pathway surveys to inform climate engineering policy choices. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(3), 560–565. https://www.jstor.org/stable/26467426

Grunwald, A. (2004). The normative basis of (health) technology assessment and the role of ethical expertise. *Poiesis & Praxis*, *2*(2), 175–193. https://doi.org/10.1007/s10202-003-0050-5

Grunwald, A. (2014). Technology Assessment for Responsible Innovation. In J. van den Hoven, N. Doorn, T. Swierstra, B.-J. Koops, & H. Romijn (Eds.), *Responsible Innovation 1: Innovative Solutions for Global Issues* (pp. 15–31). Springer Netherlands. https://doi.org/10.1007/978-94-017-8956-1_2

Grunwald, A. (2022). *Technikfolgenabschätzung: Einführung*. Nomos Verlag.

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=5272391

Grunwald, A., & Hillerbrand, R. (2021). *Handbuch Technikethik*. J. B. Metzler. https://link.springer.com/book/10.1007/978-3-476-04901-8

Guest, G., MacQueen, K. M., & Namey, E. E. (2012). *Applied Thematic Analysis*. SAGE Publications.

Guston, D. H. (2014). Understanding 'anticipatory governance.' *Social Studies of Science*, *44*(2), 218–242. https://doi.org/10.1177/0306312713508669

Guston, D. H., & Sarewitz, D. (2002). Real-time technology assessment. *Technology in Society*, *24*(1), 93–109. https://doi.org/10.1016/S0160-791X(01)00047-1

Hänsel, M. C., Drupp, M. A., Johansson, D. J. A., Nesje, F., Azar, C., Freeman, M. C., Groom, B., & Sterner, T. (2020). Climate economics support for the UN climate targets. *Nature Climate Change*, *10*(8), Article 8. https://doi.org/10.1038/s41558-020-0833-x

Harman, G. H. (1965). The Inference to the Best Explanation. *The Philosophical Review*, *74*(1), 88–95. https://doi.org/10.2307/2183532

Haszeldine, R. S., Flude, S., Johnson, G., & Scott, V. (2018). Negative emissions technologies and carbon capture and storage to achieve the Paris Agreement commitments. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *376*(2119), 20160447. https://doi.org/10.1098/rsta.2016.0447

Heyward, C. (2013). Situating and Abandoning Geoengineering: A Typology of Five Responses to Dangerous Climate Change. *PS: Political Science & Politics*, *46*(1), 23–27. https://doi.org/10.1017/S1049096512001436

Höfele, P. (2020). New technologies and the 'heuristics of fear'. The meaning and prehistory of an emotion in Jonas, Heidegger and Hegel. *Hungarian Philosophical Review*, *64*, 166–182. http://filozofiaiszemle.net/2020/12/hungarian-philosophical-review-20201-

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=5272391

self-narrativity-emotions/

Höfele, P., Reuter, L., Estadieu, L., Livanec, S., Stumpf, M., & Kiesel, A. (2022). Connecting the methods of psychology and philosophy: Applying Cognitive-Affective Maps (CAMs) to identify ethical principles underlying the evaluation of bioinspired technologies. *Philosophical Psychology*, *0*(0), 1–24. https://doi.org/10.1080/09515089.2022.2113770

Hussain, Z., Binz, M., Mata, R., & Wulff, D. U. (2024). A tutorial on open-source large language models for behavioral science. *Behavior Research Methods*. https://doi.org/10.3758/s13428-024-02455-8

Hutto, C., & Gilbert, E. (2014). VADER: A Parsimonious Rule-Based Model for Sentiment Analysis of Social Media Text. *Proceedings of the International AAAI Conference on Web and Social Media*, *8*(1), Article 1. https://doi.org/10.1609/icwsm.v8i1.14550

Jacobs, J. (2002). *Dimensions of Moral Theory: An Introduction to Metaethics and Moral Psychology*. Blackwell Publishers.

Johansson, D. J. A., Azar, C., Lehtveer, M., & Peters, G. P. (2020). The role of negative carbon emissions in reaching the Paris climate targets: The impact of target formulation in integrated assessment models. *Environmental Research Letters*, *15*(12), 124024. https://doi.org/10.1088/1748-9326/abc3f0

Johnson, R. B., & Onwuegbuzie, A. J. (2004). Mixed Methods Research: A Research Paradigm Whose Time Has Come. *Educational Researcher*, *33*(7), 14–26. https://doi.org/10.3102/0013189X033007014

Jonas, H. (2020). *Das Prinzip Verantwortung: Versuch einer Ethik für die technologische Zivilisation*. Suhrkamp Verlag.

Keller, D. P., Feng, E. Y., & Oschlies, A. (2014). Potential climate engineering effectiveness and side effects during a high carbon dioxide-emission scenario. *Nature*

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=5272391

Communications, *5*(1), Article 1. https://doi.org/10.1038/ncomms4304

Keller, K., Helgeson, C., & Srikrishnan, V. (2021). Climate Risk Management. *Annual Review of Earth and Planetary Sciences*, *49*(1), 95–116. https://doi.org/10.1146/annurev-earth-080320-055847

Klepper, G., & Rickels, W. (2012). The Real Economics of Climate Engineering. *Economics Research International*, *2012*, e316564. https://doi.org/10.1155/2012/316564

Kluyver, T., Ragan-Kelley, B., Pérez, F., Granger, B., Bussonnier, M., Frederic, J., Kelley, K., Hamrick, J., Grout, J., Corlay, S., Ivanov, P., Avila, D., Abdalla, S., Willing, C., & Jupyter development team. (2016). *Jupyter Notebooks – a publishing format for reproducible computational workflows* (F. Loizides & B. Scmidt, Eds.; pp. 87–90). IOS Press. https://doi.org/10.3233/978-1-61499-649-1-87

Kuckartz, U., & Rädiker, S. (2022). *Qualitative Inhaltsanalyse. Methoden, Praxis, Computerunterstützung*. Beltz Juventa. https://content-select.com/de/portal/media/view/5e623532-20b8-4f33-b19e-4a1db0dd2d03?forceauth=1

Le Quéré, C., Peters, G. P., Friedlingstein, P., Andrew, R. M., Canadell, J. G., Davis, S. J., Jackson, R. B., & Jones, M. W. (2021). Fossil CO2 emissions in the post-COVID-19 era. *Nature Climate Change*, *11*(3), Article 3. https://doi.org/10.1038/s41558-021-01001-0

Lee, H., & Romero, J. (2023). *IPCC, 2023: Climate Change 2023: Synthesis Report. A Report of the Intergovernmental Panel on Climate Change. Contribution of Working Groups I, II and III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*. IPCC, Geneva, Switzerland. https://www.ipcc.ch/report/ar6/syr/downloads/report/IPCC_AR6_SYR_LongerReport.pdf

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=5272391

Lichtenstein, S., & Slovic, P. (2006). *The Construction of Preference*. Cambridge University Press.

Liu, P., Yuan, W., Fu, J., Jiang, Z., Hayashi, H., & Neubig, G. (2023). Pre-train, Prompt, and Predict: A Systematic Survey of Prompting Methods in Natural Language Processing. *ACM Computing Surveys*, *55*(9), 195:1-195:35. https://doi.org/10.1145/3560815

Low, S., Boettcher, M., Asayama, S., Baum, C., Borth, A., Brown, C., Clingerman, F., Dauvergne, P., De Pryck, K., Gupta, A., Honegger, M., Lenzi, D., Reitsma, R., Schenuit, F., Scott-Buechler, C., & Valenzuela, J. M. (2024). An earth system governance research agenda for carbon removal. *Earth System Governance*, *19*, 100204. https://doi.org/10.1016/j.esg.2024.100204

Lucivero, F. (2016). *Ethical Assessments of Emerging Technologies: Appraising the moral plausibility of technological visions*. Springer.

MacMartin, D. G., Irvine, P. J., Kravitz, B., & Horton, J. B. (2019). Technical characteristics of a solar geoengineering deployment and implications for governance. *Climate Policy*, *19*(10), 1325–1339. https://doi.org/10.1080/14693062.2019.1668347

Macnaghten, P., Davies, S. R., & Kearnes, M. (2019). Understanding Public Responses to Emerging Technologies: A Narrative Approach. *Journal of Environmental Policy & Planning*, *21*(5), 504–518. https://doi.org/10.1080/1523908X.2015.1053110

Marchau, V. A. W. J., Walker, W. E., Bloemen, P. J. T. M., & Popper, S. W. (Eds.). (2019). *Decision Making under Deep Uncertainty*. Springer Nature. https://doi.org/10.1007/978-3-030-05252-2

Masson-Delmotte, V., Zhai, P., Pörtner, H.-O., Roberts, D., Skea, J., Shukla, P. R., Pirani, A., Moufouma-Okia, W., Péan, C., Pidcock, R., Connors, S., Matthews, J. B. R., Chen, Y., Zhou, X., Gomis, M. I., Lonnoy, E., Maycock, T., Tignor, M., & Waterfield, T. (2018). *IPCC, 2018: Global Warming of 1.5°C. An IPCC Special Report on the*

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=5272391

impacts of global warming of 1.5°C above pre-industrial levels and related global greenhouse gas emission pathways, in the context of strengthening the global response to the threat of climate change, sustainable development, and efforts to eradicate poverty. Cambridge University Press. https://www.ipcc.ch/sr15/

Mayring, P. (2022). *Qualitative Content Analysis: A Step-by-Step Guide*. SAGE.

McCain, K., & Poston, T. (2017). *Best Explanations: New Essays on Inference to the Best Explanation*. Oxford University Press.

McLaren, D., Parkhill, K. A., Corner, A., Vaughan, N. E., & Pidgeon, N. F. (2016). Public conceptions of justice in climate engineering: Evidence from secondary analysis of public deliberation. *Global Environmental Change*, *41*, 64–73. https://doi.org/10.1016/j.gloenvcha.2016.09.002

McMillan, J. (2018). *The methods of bioethics: An essay in meta-bioethics* (First edition). Oxford University Press.

Meadows, D. H., Meadows, D. L., Randers, J., & Iii, W. W. B. (1972). *The Limits to Growth—Club of Rome*. Club of Rome. https://www.clubofrome.org/publication/the-limits-to-growth/

Merk, C., Pönitzsch, G., Kniebes, C., Rehdanz, K., & Schmidt, U. (2015). Exploring public perceptions of stratospheric sulfate injection. *Climatic Change*, *130*(2), 299–312. https://doi.org/10.1007/s10584-014-1317-7

Mirzadeh, I., Alizadeh, K., Shahrokhi, H., Tuzel, O., Bengio, S., & Farajtabar, M. (2024). *GSM-Symbolic: Understanding the Limitations of Mathematical Reasoning in Large Language Models* (No. arXiv:2410.05229). arXiv. https://doi.org/10.48550/arXiv.2410.05229

Möller, M., & Grießhammer, R. (2022). Prospective technology assessment in the Anthropocene: A transition toward a culture of sustainability. *The Anthropocene*

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=5272391

Review, *9*(2), 257–275. https://doi.org/10.1177/20530196221095700

Moosbrugger, H., & Kelava, A. (Eds.). (2020). *Testtheorie und Fragebogenkonstruktion*. Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-662-61532-4

Musschenga, B. (2009). Was ist empirische Ethik? *Ethik in der Medizin*, *21*(3), 187–199. https://doi.org/10.1007/s00481-009-0025-8

Muthén, L. K., & Muthen, B. O. (2017). *Mplus User's Guide. Eighth Edition.* [Computer software]. Muthén & Muthén. https://www.statmodel.com/html_ug.shtml

National Research Council. (2015). *Climate Intervention: Reflecting Sunlight to Cool Earth*. National Academies Press.

Neuber, F. (2018). *Buying Time with Climate Engineering? An analysis of the buying time framing in favor of climate engineering* [PhD Thesis, Karlsruher Institut für Technologie (KIT)]. https://doi.org/10.5445/IR/1000084294

Neuhäuser, C., Raters, M.-L., & Stoecker, R. (Eds.). (2023). *Handbuch Angewandte Ethik*. J.B. Metzler. https://doi.org/10.1007/978-3-476-05869-0

Newman, M. (2018). *Networks: An Introduction*. Oxford University Press.

Nordhaus, W. (1992). The 'Dice' Model: Background and Structure of a Dynamic Integrated Climate-Economy Model of the Economics of Global Warming. *Cowles Foundation Discussion Papers*. https://elischolar.library.yale.edu/cowles-discussion-paper-series/1252

Nordhaus, W. (2018). Projections and Uncertainties about Climate Change in an Era of Minimal Climate Policies. *American Economic Journal: Economic Policy*, *10*(3), 333–360. https://doi.org/10.1257/pol.20170046

Ott, K. (2011). Argumente für und wider „Climate Engineering ". In *Fallstudien zur Ethik in Wissenschaft, Wirtschaft, Technik und Gesellschaft* (pp. 198–210). KIT Scientific Publishing.

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=5272391

Ott, K. (2012). Domains of Climate Ethics. *Jahrbuch für Wissenschaft und Ethik*, *16*(1), 95–114. https://doi.org/10.1515/jfwe.2012.95

Ott, K., & Neuber, F. (2020). Climate engineering. In *Oxford Research Encyclopedia of Climate Science*. Oxford University Press.

Ouyang, L., Wu, J., Jiang, X., Almeida, D., Wainwright, C. L., Mishkin, P., Zhang, C., Agarwal, S., Slama, K., Ray, A., Schulman, J., Hilton, J., Kelton, F., Miller, L., Simens, M., Askell, A., Welinder, P., Christiano, P., Leike, J., & Lowe, R. (2022). *Training language models to follow instructions with human feedback* (No. arXiv:2203.02155). arXiv. https://doi.org/10.48550/arXiv.2203.02155

Pachauri, R. K., & Meyer, L. A. (2014). *IPCC, 2014: Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. https://www.ipcc.ch/report/ar5/syr/

Palm, E., & Hansson, S. O. (2006). The case for ethical technology assessment (eTA). *Technological Forecasting and Social Change*, *73*(5), 543–558. https://doi.org/10.1016/j.techfore.2005.06.002

Parkhill, K., & Pidgeon, N. (2011). Public Engagement on Geoengineering Research: Preliminary Report on the SPICE Deliberative Workshops. In *Public Engagement on Geoengineering Research*. https://eprints.whiterose.ac.uk/82892/

Parkhill, K., Pidgeon, N., Corner, A., & Vaughan, N. (2013). Deliberation and Responsible Innovation: A Geoengineering Case Study. In *Responsible Innovation* (pp. 219–239). John Wiley & Sons, Ltd. https://doi.org/10.1002/9781118551424.ch12

Paulo, N., & Bublitz, J. C. (2020). *Empirische Ethik: Grundlagentexte aus Psychologie und Philosophie*. Suhrkamp Verlag.

Pidgeon, N. (2021). Engaging publics about environmental and technology risks: Frames, values and deliberation. *Journal of Risk Research*, *24*(1), 28–46.

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=5272391

https://doi.org/10.1080/13669877.2020.1749118

Pidgeon, N., Parkhill, K., Corner, A., & Vaughan, N. (2013). Deliberating stratospheric aerosols for climate geoengineering and the SPICE project. *Nature Climate Change*, *3*(5), Article 5. https://doi.org/10.1038/nclimate1807

Pieper, A. (2017). *Einführung in die Ethik*. UTB.

Plazzotta, M., Séférian, R., Douville, H., Kravitz, B., & Tjiputra, J. (2018). Land Surface Cooling Induced by Sulfate Geoengineering Constrained by Major Volcanic Eruptions. *Geophysical Research Letters*, *45*(11), 5663–5671. https://doi.org/10.1029/2018GL077583

Pokorny, J. J., Norman, A., Zanesco, A. P., Bauer-Wu, S., Sahdra, B. K., & Saron, C. D. (2018). Network analysis for the visualization and analysis of qualitative data. *Psychological Methods*, *23*(1), 169–183. https://doi.org/10.1037/met0000129

Pörtner, H.-O., Roberts, D. C., Tignor, M., Poloczanska, E. S., Mintenbeck, K., Alegría, A., Craig, M., Langsdorf, S., Löschke, S., Möller, V., Okem, A., & Rama, B. (2022). *IPCC, 2022: Climate Change 2022: Impacts, Adaptation, and Vulnerability. Contribution of Working Group II to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press. https://www.ipcc.ch/report/sixth-assessment-report-working-group-ii/

Preston, C. J. (2012). *Engineering the Climate*. Rowman & Littlefield.

Preston, C. J. (2013). Ethics and geoengineering: Reviewing the moral issues raised by solar radiation management and carbon dioxide removal. *WIREs Climate Change*, *4*(1), 23–37. https://doi.org/10.1002/wcc.198

R Core Team. (2020). *R: A Language and Environment for Statistical Computing* [Computer software]. R Foundation for Statistical Computing. https://www.R-project.org/

Raschka, S. (2024). *Build a Large Language Model (From Scratch)*. Simon and Schuster.

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=5272391

Reuter, L., Mansell, J., Rhea, C., & Kiesel, A. (2022). Direct assessment of individual connotation and experience: An introduction to cognitive-affective mapping. *Politics and the Life Sciences*, *41*(1), 131–139. https://doi.org/10.1017/pls.2021.31

Reynolds, J. L., & Horton, J. B. (2020). An earth system governance perspective on solar geoengineering. *Earth System Governance*, *3*, 100043. https://doi.org/10.1016/j.esg.2020.100043

Rickels, W., Klepper, G., Dovern, J., Betz, G., Nadine, B., Güssow, K., Heintzenberg, J., Hiller, S., Hoose, C., Leisner, T., Oschlies, A., Platt, U., Proelß, A., Schäfer, S., Zürn, M., Cacean, S., & Renn, O. (2011). *Large-Scale Intentional Intervention s into the Climate System? Assessing the Climate Engineering Debate* [Scoping Report]. https://www.ifw-kiel.de/de/publikationen/books/large-scale-intentional-intervention-s-into-the-climate-system-assessing-the-climate-engineering-debate-6632/

Rockström, J., Steffen, W., Noone, K., Persson, Å., Chapin, F. S., Lambin, E. F., Lenton, T. M., Scheffer, M., Folke, C., Schellnhuber, H. J., Nykvist, B., de Wit, C. A., Hughes, T., van der Leeuw, S., Rodhe, H., Sörlin, S., Snyder, P. K., Costanza, R., Svedin, U., … Foley, J. A. (2009). A safe operating space for humanity. *Nature*, *461*(7263), Article 7263. https://doi.org/10.1038/461472a

Sand, M., Hofbauer, B. P., & Alleblas, J. (2023). Techno-fixing non-compliance—Geoengineering, ideal theory and residual responsibility. *Technology in Society*, *73*, 1–9. https://doi.org/10.1016/j.techsoc.2023.102236

Schwartz, M. S. (2016). Ethical Decision-Making Theory: An Integrated Approach. *Journal of Business Ethics*, *139*(4), 755–776. https://doi.org/10.1007/s10551-015-2886-8

Shepherd, J. G. (2009). *Geoengineering the climate: Science, governance and uncertainty* (p. 98) [Monograph]. Royal Society. https://doi.org/10/29)

Siegrist, M., & Árvai, J. (2020). Risk Perception: Reflections on 40 Years of Research. *Risk*

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=5272391

Analysis, *40*(S1), 2191–2206. https://doi.org/10.1111/risa.13599

Siegrist, M., & Hartmann, C. (2020). Consumer acceptance of novel food technologies. *Nature Food*, *1*(6), Article 6. https://doi.org/10.1038/s43016-020-0094-x

Sonntag, S., Ferrer González, M., Ilyina, T., Kracher, D., Nabel, J. E. M. S., Niemeier, U., Pongratz, J., Reick, C. H., & Schmidt, H. (2018). Quantifying and Comparing Effects of Climate Engineering Methods on the Earth System. *Earth's Future*, *6*(2), 149–168. https://doi.org/10.1002/2017EF000620

Sovacool, B. K., Baum, C. M., & Low, S. (2022). Determining our climate policy future: Expert opinions about negative emissions and solar radiation management pathways. *Mitigation and Adaptation Strategies for Global Change*, *27*(8), 58. https://doi.org/10.1007/s11027-022-10030-9

Sovacool, B. K., Baum, C. M., & Low, S. (2023). Beyond climate stabilization: Exploring the perceived sociotechnical co-impacts of carbon removal and solar geoengineering. *Ecological Economics*, *204*, 107648. https://doi.org/10.1016/j.ecolecon.2022.107648

Steegen, S., Tuerlinckx, F., Gelman, A., & Vanpaemel, W. (2016). Increasing Transparency Through a Multiverse Analysis. *Perspectives on Psychological Science*, *11*(5), 702–712. https://doi.org/10.1177/1745691616658637

Sühr, T., Dorner, F. E., Samadi, S., & Kelava, A. (2024). *Challenging the Validity of Personality Tests for Large Language Models* (No. arXiv:2311.05297). arXiv. https://doi.org/10.48550/arXiv.2311.05297

ter Mors, E., Terwel, B. W., Daamen, D. D. L., Reiner, D. M., Schumann, D., Anghel, S., Boulouta, I., Cismaru, D. M., Constantin, C., de Jager, C. C. H., Dudu, A., Esken, A., Falup, O. C., Firth, R. M., Gemeni, V., Hendriks, C., Ivan, L., Koukouzas, N., Markos, A., … Ziogou, F. (2013). A comparison of techniques used to collect informed public opinions about CCS: Opinion quality after focus group discussions

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=5272391

versus information-choice questionnaires. *International Journal of Greenhouse Gas Control*, *18*, 256–263. https://doi.org/10.1016/j.ijggc.2013.07.015

Thagard, P. (1998). Ethical coherence. *Philosophical Psychology*, *11*(4), 405–422. https://doi.org/10.1080/09515089808573270

Thagard, P. (2000). *Coherence in Thought and Action*. MIT Press.

Thagard, P. (2010). EMPATHICA: A Computer Support System with Visual Representations for Cognitive-Affective Mapping. *Workshops at the Twenty-Fourth AAAI Conference on Artificial Intelligence*, 79–81. https://www.aaai.org/ocs/index.php/WS/AAAIW10/paper/view/1981

Thomas, G., Pidgeon, N., & Roberts, E. (2018). Ambivalence, naturalness and normality in public perceptions of carbon capture and storage in biomass, fossil energy, and industrial applications in the United Kingdom. *Energy Research & Social Science*, *46*, 1–9. https://doi.org/10.1016/j.erss.2018.06.007

Touvron, H., Lavril, T., Izacard, G., Martinet, X., Lachaux, M.-A., Lacroix, T., Rozière, B., Goyal, N., Hambro, E., Azhar, F., Rodriguez, A., Joulin, A., Grave, E., & Lample, G. (2023). *LLaMA: Open and Efficient Foundation Language Models* (No. arXiv:2302.13971). arXiv. https://doi.org/10.48550/arXiv.2302.13971

Tunstall, L., Werra, L. von, & Wolf, T. (2022). *Natural Language Processing with Transformers*. O'Reilly Media, Inc.

United Nations Environment Programme. (1992). *Report of the United Nations Conference on Environment and Development*. https://www.un.org/en/conferences/environment/rio1992

Van Rossum, G., & Drake, F. L. (2009). *Python 3 Reference Manual*. CreateSpace.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł. ukasz, & Polosukhin, I. (2017). Attention is All you Need. *Advances in Neural*

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=5272391

Information Processing Systems, *30*. https://proceedings.neurips.cc/paper/2017/hash/3f5ee243547dee91fbd053c1c4a845aa-Abstract.html

Vink, S., van Tartwijk, J., Verloop, N., Gosselink, M., Driessen, E., & Bolk, J. (2016). The articulation of integration of clinical and basic sciences in concept maps: Differences between experienced and resident groups. *Advances in Health Sciences Education*, *21*(3), 643–657. https://doi.org/10.1007/s10459-015-9657-2

Wang, Y., Ma, X., Zhang, G., Ni, Y., Chandra, A., Guo, S., Ren, W., Arulraj, A., He, X., Jiang, Z., Li, T., Ku, M., Wang, K., Zhuang, A., Fan, R., Yue, X., & Chen, W. (2024). *MMLU-Pro: A More Robust and Challenging Multi-Task Language Understanding Benchmark (Published at NeurIPS 2024 Track Datasets and Benchmarks)* (No. arXiv:2406.01574). arXiv. https://doi.org/10.48550/arXiv.2406.01574

Welsby, D., Price, J., Pye, S., & Ekins, P. (2021). Unextractable fossil fuels in a 1.5 °C world. *Nature*, *597*(7875), Article 7875. https://doi.org/10.1038/s41586-021-03821-8

White, J., Fu, Q., Hays, S., Sandborn, M., Olea, C., Gilbert, H., Elnashar, A., Spencer-Smith, J., & Schmidt, D. C. (2023). *A Prompt Pattern Catalog to Enhance Prompt Engineering with ChatGPT* (No. arXiv:2302.11382). arXiv. https://doi.org/10.48550/arXiv.2302.11382

Wibeck, V., Hansson, A., Anshelm, J., Asayama, S., Dilling, L., Feetham, P. M., Hauser, R., Ishii, A., & Sugiyama, M. (2017). Making sense of climate engineering: A focus group study of lay publics in four countries. *Climatic Change*, *145*(1), 1–14. https://doi.org/10.1007/s10584-017-2067-0

Workman, M., Dooley, K., Lomax, G., Maltby, J., & Darch, G. (2020). Decision making in contexts of deep uncertainty—An alternative approach for long-term climate policy. *Environmental Science & Policy*, *103*, 77–84.

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=5272391

https://doi.org/10.1016/j.envsci.2019.10.002

Xie, Y., Allaire, J. J., & Grolemund, G. (2018). *R Markdown: The Definitive Guide*. Chapman and Hall/CRC. https://doi.org/10.1201/9781138359444

Yan, L., Sha, L., Zhao, L., Li, Y., Martinez-Maldonado, R., Chen, G., Li, X., Jin, Y., & Gašević, D. (2024). Practical and ethical challenges of large language models in education: A systematic scoping review. *British Journal of Educational Technology*, *55*(1), 90–112. https://doi.org/10.1111/bjet.13370

Yang, J., Jin, H., Tang, R., Han, X., Feng, Q., Jiang, H., Zhong, S., Yin, B., & Hu, X. (2024). Harnessing the Power of LLMs in Practice: A Survey on ChatGPT and Beyond. *ACM Trans. Knowl. Discov. Data*, *18*(6), 160:1-160:32. https://doi.org/10.1145/3649506

Zhang, H., Wang, F., Li, J., Duan, Y., Zhu, C., & He, J. (2022). Potential Impact of Tonga Volcano Eruption on Global Mean Surface Air Temperature. *Journal of Meteorological Research*, *36*(1), 1–5. https://doi.org/10.1007/s13351-022-2013-6

Declaration of generative AI and AI-assisted technologies in the writing process

During the preparation of this work, the authors utilized one Llama model, which were developed by Meta to analyze data and generate textual summaries. Specifically, the AI model "Llama-3.1-70B-Instruct" was employed for generating structured summaries and comparative syntheses of qualitative data based on predefined prompts. These prompts guided the model to extract and organize key ethical arguments from coded datasets, and to compare layperson interpretations with expert definitions. After using this tool, the authors thoroughly reviewed and edited the content to ensure accuracy, coherence, and alignment with the study's objectives. The authors take full responsibility for the final content of the published article.

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=5272391

Appendix A: Sample statistics Cognitive-Affective Maps

The subsequent report was generated by utilizing the “Get Report” function within the Data Analysis Tool (Fenn et al., under review).

Description of dataset

In total, we collected 58 CAMs, of which 0 (0%) CAMs were excluded from further analysis. Participants drew on average 25.4 (*SD* = 2.06) concepts (whereby 34% were positive, 46% negative, 12% neutral and 8% ambivalent). Please note that the technical settings required participants to draw at least 24 concepts. On average, 44.21 (*SD* = 32.49) connectors were drawn. 82% of the connectors were agreeing and 18% disagreeing. Furthermore, 21% of the connectors were bidirectional and 79% unidirectional. The valence for the concepts range from [-3,-1] for negative and [1,3] for positive concepts, with ambivalent and neutral concepts being assigned a value of 0. The mean average valence over all the CAMs was -0.33 (*SD* = 0.51). In 14% of the non-deleted CAMs one or more of the predefined concepts were removed by the participants.

Summarizing concepts

We summarized the CAMs using the dedicated Data Analysis Tool. The Data Analysis Tool generates a protocol which tracks each summarizing step so that the summarizing process is completely transparent. The 1,063 raw unique concepts (1,473 in total) were summarized to 41 concepts using 101 times the “approximate matching”, 203 times the “searching terms”, 4 times the “search for synonyms” and 10 times the “apply word2vec model” functionalities.

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=5272391

Appendix B: Summary of Ethical Arguments in Cognitive-Affective Maps

In table SM2, the frequencies and average valence of all summarized concepts to ethical arguments are presented (cf., Tab. 1), whereby an overall searchable wordlist can be found on OSF (https://osf.io/rhxfn). In the table, the last row "mimics nature" is not an ethical argument but could indicate the above discussed "Natural-is-better" heuristic (see section 2.1 in main article). For the meaning of single variables please see "Note" below the table.

Table SM2.
Frequencies of all summarized concepts to ethical arguments

Concept	*N*	*M*	*SD*	Examples
Side-effects not predictable	264	-2.25	0.98	uncertain whether it would contribute to more acid rain; unknown side effects
Greater Good	143	2.22	0.9	environmental benefits; improve global health
Risk of Governance	63	-1.86	1.05	government focuses policy on electoral cycles [..]
Moral Hazard	37	-1.46	1.76	allows society to carry on polluting the planet
Risk Transfer to the Future	37	-2.32	0.82	social or political conflicts over the use; cause war?
Informed Consent	31	-1.32	1.19	impossible to get everyone to agree on it
Long-Term Control	24	-1.92	0.97	could be hard to get all countries to commit for a long period of time
Technological Fix	21	0.48	1.97	a good solution [..]; less human burden
Termination Problem	19	-1.84	1.17	if stopped temperatures would rise rapidly
Unfair distribution of effects and power	17	-1.41	1.12	might affect one region more than another, causing tensions
Arming the Future	14	2	1.11	creating a better future for future generations
Maintaining the Status Quo	12	-1.83	1.11	companies, governments will carry on polluting and using bad technology/fossil fuels
Buying Time	11	1	0.77	allows for more time to create long term solutions
Lesser-evil	8	-0.62	1.19	emergency solution
Risk of Unilateral Use	8	-2.75	0.46	governments in wealthier nations would be able to exert control over poorer nations
Betrayal of Divine Creation	6	-2.33	0.82	Is SAI another way in which humans mess things up by playing God?
Do it Alone	4	1.5	1.29	can be implemented by a wide variety of nations
Dual Use	3	-3	0	governments could weaponise [..] the use
Hubris Argument	2	-1	1.41	give us the illusion we can carry on as we have been
mimics nature	10	1.3	1.16	Technology that comes from [..] volcanos

Note. N is the total frequency of summarized concepts to ethical argument, *M* the mean valence, *SD* the standard deviation of the mean valence, and in the Examples column are typical examples from the text or comments of the drawn concepts.

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=5272391

Appendix C: Summary of Assigned Ethical Arguments in Open Text

In table SM3, the frequencies and percentages of all assigned ethical arguments are presented. In total, only 553 ethical arguments were considered (see main article). Raters could assign multiple ethical arguments to the text answer of a single participant (for meaning of single variables, see “Note”).

Table SM3.
Frequencies of all assigned ethical arguments

Ethical argument	*N*	percentage	number of co-occurrence
Emergency Case	202	23.01	160
Informed Consent	126	14.35	153
Lesser-evil	119	13.55	94
99 (residual category)	87	9.91	87
Side-effects unseen / not predictable	74	8.43	83
Greater Good	71	8.09	38
Technological Fix	31	3.53	36
Unfair distribution of effects and power	26	2.96	50
Moral Hazard	24	2.73	33
Termination Problem	18	2.05	30
Arming the Future	17	1.94	14
Hubris Argument	15	1.71	16
Buying Time	14	1.59	18
Risk of Unilateral Use	12	1.37	21
Maintaining the Status Quo	10	1.14	16
Risk of Governance	10	1.14	17
Betrayal of Divine Creation	7	0.80	11
Long-Term Control	6	0.68	12
Risk Transfer to the Future	4	0.46	8
Do it Alone	3	0.34	4
Unstoppable Deployment if researched	2	0.23	7
Dual Use	0	0	0

Note. *N* is the frequency a single ethical argument was assigned by all raters, percentage is the respective percentage (divided by total sum of number) and the “number of co-occurrence” indicates how often a single ethical argument was assigned together with all the other ethical arguments.

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=5272391

Appendix D: Introduction to Large Language Models

LLMs are advanced artificial intelligence tools that excel at processing and generating human-like text, which is shown by benchmark testing (e.g., Chiang et al., 2024; Wang et al., 2024). These models are trained on extensive datasets comprising billions to trillions of tokens - fundamental units of text processed by LLMs, which may represent entire words, subwords, or characters, depending on the tokenizer - enabling them to identify patterns, relationships, and structures inherent in language (Caelen & Blete, 2023; Tunstall et al., 2022). A defining characteristic of LLMs is their ability to predict the next word or token in a sequence. For example, when prompted with "The capital of France is", an LLM will likely predict "Paris," leveraging probabilistic patterns learned from its training data (Hussain et al., 2024; Vaswani et al., 2017)

LLMs are built on the generative pretrained transformer (GPT) architecture, which relies on self-attention mechanisms to effectively process input text. Self-attention enables the model to focus on the most relevant parts of the input sequence, capturing contextual meaning at varying scales (Vaswani et al., 2017). This architecture excels across a wide array of tasks, including summarization, machine translation, classification, and creative writing (Dubey et al., 2024; Hussain et al., 2024; Touvron et al., 2023). The training of an LLM typically involves two stages: pre-training and fine-tuning. During pre-training, the model is exposed to a vast corpora, including books, scientific articles, and internet-sourced text, to learn general language patterns resulting in a foundational model (Raschka, 2024). Fine-tuning subsequently adapts the pre-trained model to specific tasks or domains using curated datasets and human feedback (Brown et al., 2020; Ouyang et al., 2022). Fine-tuned models often include the term "Instruct" in their name to denote alignment with task-specific objectives, as seen in the Llama-3.1-70B-Instruct model (Dubey et al., 2024). LLMs are versatile tools with applications spanning multiple domains. They can condense large datasets

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=5272391

into concise, interpretable summaries, classify text into predefined categories, generate synthetic data for machine learning applications, and assist researchers in identifying themes and patterns within complex datasets (Debelak et al., 2024; Hussain et al., 2024; Yang et al., 2024).

Despite their transformative potential, LLMs have inherent limitations. They excel at pattern recognition and text generation but lack genuine understanding, reasoning, or cognitive abilities akin to humans (Mirzadeh et al., 2024). Their reliance on training data makes them susceptible to perpetuating biases or inaccuracies present in the underlying datasets (Sühr et al., 2024; Yan et al., 2024). Furthermore, their context length - the number of tokens that can be processed simultaneously - varies between models, ranging from 8,000 tokens to over 100,000 tokens, which can restrict their utility for lengthy or complex tasks (Dubey et al., 2024; Tunstall et al., 2022).

Effective use of LLMs requires adherence to well-established prompting strategies. Structured prompts, which include clear instructions, relevant context, and examples of desired outputs, can significantly enhance model performance (Liu et al., 2023; White et al., 2023).

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=5272391